

Effectiveness of AI-Assisted Orthopedic Nursing Education: A Mixed-Methods Randomized Controlled

Zhixin Yao

Third Affiliated Hospital of Southern Medical University

Gan lin

Third Affiliated Hospital of Southern Medical University

Yiyuan Li

Third Affiliated Hospital of Southern Medical University

Sitong He

Third Affiliated Hospital of Southern Medical University

Yongkang Chen

Third Affiliated Hospital of Southern Medical University

Jingjing Pan

Third Affiliated Hospital of Southern Medical University

Guangyuan Dong

Third Affiliated Hospital of Southern Medical University

Hui Ouyang

Third Affiliated Hospital of Southern Medical University

Yating Li

Third Affiliated Hospital of Southern Medical University

Fang Qin

Southern Medical University

Jinghui Yao

Third Affiliated Hospital of Southern Medical University

Yun Chen

chenyun88@smu.edu.cn

Third Affiliated Hospital of Southern Medical University

Research Article

Keywords: Orthopedic nursing, Nursing education, Artificial intelligence, Image assessment, Nursing decision-making

Posted Date: August 18th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10274926/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Orthopedic nursing education requires learners to convert imaging cues into patient assessment, risk identification, and care planning, but traditional teaching provides limited visualization and feedback. This study aimed to evaluate and interpret the effectiveness of an AI-assisted orthopedic nursing education using DDH as a teaching case.

Methods

Sixty nursing interns were randomized, and 56 participants with complete post-training data were included in the final analysis. Both groups received the same DDH educational content, and the intervention group additionally used an AI-assisted tool. We analyzed the quantitative data using IBM SPSS Statistics 27.0, including chi-square or Fisher's exact tests, independent-samples t tests, repeated-measures analysis of variance, Mann–Whitney U tests, and Wilcoxon signed-rank tests. Interviews with 10 intervention participants were analyzed using Colaizzi's method.

Results

The intervention group had higher post-training scores on DDH-related nursing assessment (MD = 3.72, 95% CI [0.96, 6.48], $d = 0.72$, $p = 0.009$), DDH classification (MD = 4.07, 95% CI [0.17, 7.97], $d = 0.56$, $p = 0.041$), and deep learning (MD = 5.64, 95% CI [0.25, 11.03], $d = 0.56$, $p = 0.040$). However, the time-by-group interactions for deep learning and DDH-related nursing assessment were not statistically significant, so the evidence for differential pre-post change between groups was limited. In exploratory analyses, several AI learning intention factors showed within-group increases in the intervention group. Qualitative findings indicated that AI-assisted imaging education helped interns integrate image interpretation with patient assessment, nursing decision-making, and care planning. The findings also highlighted potential limitations, including the risk of undermining independent thinking and insufficient consideration of students' baseline knowledge and abilities.

Conclusions

AI-assisted orthopedic nursing education was associated with higher post-training scores in DDH classification and nursing assessment, by helping interns translate imaging findings into clinically relevant information. However, several secondary outcomes did not show significant group-by-time interaction effects. Therefore, these findings should be interpreted as preliminary. Future implementation should preserve opportunities for independent thinking and incorporate differentiated instruction tailored to students' baseline knowledge and abilities.

Trial registration

The study was prospectively registered with the Chinese Clinical Trial Registry (ChiCTR2600119719) on March 3, 2026.

1. Introduction

Orthopedic nursing influences patients' functional recovery and quality of life across the continuum from injury to rehabilitation [1]. Imaging education serves as a vital link between anatomical structures and clinical assessment [2]. Although nurses are not responsible for making final diagnostic interpretations of imaging findings, their understanding of this information helps them assess the severity of a patient's condition, identify nursing risks, explain treatment plans, and provide a basis for positioning management, rehabilitation guidance, and health education [3]. However, traditional orthopedic nursing education still presents several barriers to imaging-related learning. Static textbook and PowerPoint images make it difficult to demonstrate skeletal structures dynamically [4], while limited case resources restrict learners' exposure to varied case types and levels of difficulty [5]. In addition, feedback is often delayed because students depend on instructor comments or post-class discussion after completing image assessments[6]. Standardized, centralized teaching models also provide limited support for personalized learning pathways and targeted practice [7].

In recent years, artificial intelligence (AI) has increasingly been integrated into nursing and health professional education, demonstrating particular potential in areas such as image interpretation[6], case-based training [8], simulation training [9], and feedback support [10]. However, existing research has largely focused on quantitative learning outcomes and the effectiveness of AI applications. Learner experience, perceived barriers, trust in AI, and the risks of overreliance have not yet been fully explored [8]. In addition, existing studies have primarily evaluated image recognition accuracy and diagnostic classification, with less attention paid to nursing competencies such as nursing decision-making and health education [11].

Using developmental dysplasia of the hip (DDH) as a representative orthopedic imaging case, this study aimed to evaluate the value of AI-assisted imaging education in orthopedic nursing, with attention to learning outcomes, learner experiences, perceived barriers, trust in AI, and potential overreliance. The findings provide evidence for integrating AI-assisted imaging tools into specialized nursing education to enhance teaching quality and improve the learning experience.

2. Methods

2.1. Study Design

This study employed a mixed-methods design. The first phase was a single-center, parallel-group randomized controlled trial reported in accordance with the Consolidated Standards of Reporting Trials

(CONSORT) statement, using DDH as a teaching case to evaluate the effectiveness of AI-assisted orthopedic nursing education. The second phase was a qualitative phenomenological study that explored learners' experiences after participation in AI-assisted education. The trial was registered with the Chinese Clinical Trial Registry (ChiCTR2600119719) on March 3, 2026, before randomization and intervention implementation. The trial protocol can be obtained from this source.

No public were involved in the design, conduct, reporting, or dissemination plans of this study.

2.2. Participants and Settings

This study used convenience sampling to recruit nursing interns from the Third Affiliated Hospital of Southern Medical University. Eligible interns who met the inclusion criteria and voluntarily agreed to participate were invited to enter the formal study phase in March 2026. The inclusion criteria were as follows: (1) full-time nursing students; (2) completion of basic courses such as anatomy, histology, and surgical nursing; and (3) informed consent and voluntary participation. The exclusion criteria were as follows: (1) previous completion of a course on DDH imaging interpretation; (2) participation in similar educational research studies.

The sample size was estimated using G*Power version 3.1.9.6 with the DDH classification test score as the primary outcome [12]. The effect size was set at *Cohen's d* = 0.8 based on a previous education study with a comparable pre-post learning outcome [13], with $\alpha = 0.05$, power = 0.80, and a 1:1 allocation ratio for two groups. The required sample size was 52 participants. After allowing for a 10% attrition rate, 60 participants were recruited. Four participants (two in each group) did not complete the assigned training and had no post-training outcome data. Therefore, 56 participants were included in the final analysis.

In the qualitative research phase, purposive sampling was used to select participants from those who had participated in the first phase of the study. Interviewees were selected according to the principle of maximum variation to ensure the representation of students with diverse learning characteristics. The sample size was determined based on the principle of thematic saturation.

No interim analyses or formal stopping guidelines were planned

2.3. Procedure

2.3.1. Development of the AI Tool

Before the intervention, a multidisciplinary research team comprising members from nursing, medicine, and engineering developed the deep learning-based DDH diagnostic tool used in this study. The tool identifies six pelvic anatomical landmarks, calculates radiographic indicators relevant to DDH assessment, and classifies DDH into four categories: normal, suspected dislocation, mild dislocation, and severe dislocation. Previous validation showed high detection and classification performance, including a precision of 0.978, recall of 0.983, mAP50 of 0.993, mAP50-95 of 0.882, and overall

classification accuracy of 0.850 [14]. The tool was deployed as an open-access website (www.ddhip.com), accessible by computer, tablet, or mobile phone.

2.3.2. Allocation and Blinding

Randomization was performed after eligibility assessment and informed consent. IBM SPSS Statistics 27.0 was used to generate 60 random numbers, each linked to a participant screening number. An independent research assistant who was unaware of the study hypotheses sorted eligible participants by the generated random numbers and assigned them sequentially to the intervention group or control group in a 1:1 ratio.

Because the intervention group used the AI tool during training, participants and instructors could not be fully blinded after allocation. To reduce assessment and analysis bias, questionnaire distribution, data collection, and data entry were performed by an independent nursing postgraduate who did not participate in the teaching intervention. During data entry and analysis, the two groups were coded as “A” and “B”, and the true group identities were disclosed to the research team only after the analyses and conclusions had been completed.

2.3.3. Intervention

All participants completed baseline questionnaires and DDH-related nursing assessment before training. Both groups then entered a 17-day training program that combined a face-to-face teaching session with online case-based learning. The face-to-face session consisted of a 30-minute PowerPoint lecture delivered by a pediatric orthopedic nurse, covering DDH-related theoretical knowledge, imaging diagnosis and classification, and nursing interventions, followed by a 10-minute in-class exercise.

After the face-to-face session, online case materials adapted from real clinical scenarios were released to both groups on days 1, 2, 3, 6, 9, and 15 according to review intervals based on the Ebbinghaus forgetting curve [15]. Each release included approximately 10 minutes of online discussion and covered DDH imaging classification, treatment options, and nursing decision-making. The teaching team provided standardized feedback after each case release. The two groups had the same training schedule, lecture content, instructor, in-class exercises, online release times, case discussion format, post-training testing procedure, and assessment indicators.

The intervention group additionally used the deep learning-based DDH diagnostic tool during the lecture, in-class practice, and online case learning. The tool was demonstrated as an imaging education assistant and could be used by learners to compare their initial image judgments with AI-generated landmark identification and classification feedback. Participants in the intervention group could continue to access the tool after the course. The control group completed the same teaching and practice activities but did not have access to the AI diagnostic tool during the training period.

No additional concomitant educational support or co-interventions were provided to either group during the trial.

2.3.4. Intervention fidelity and contamination control

Before the course, all teaching assistants participated in a collective preparation session and rehearsal to standardize the lecture content, AI-tool demonstration, online case materials, answer explanations, and assessment procedures. Both groups were taught by the same pediatric orthopedic nurse to maintain consistency in core teaching content. Participants were instructed not to share course materials across groups during the intervention period to reduce contamination. All post-training tests were administered at the same time using the same assessment content and procedures, and questionnaires were completed anonymously.

2.3.5. Interview

The study adhered to the Consolidated Criteria for Reporting Qualitative Research (COREQ). The interview guide was developed based on a literature review and internal discussions within the research team. Two experts with long-term experience in nursing education and research reviewed the preliminary interview guide. After expert consultation, a pilot interview was conducted with two nursing interns who had participated in the DDH training, and the guide was then finalized. The final interview guide included the following questions: (1) How did you find the clinical training using the deep learning-based diagnostic tool for DDH, and what did you gain from the experience? (2) How has this diagnostic tool affected your clinical learning and practical skills? (3) Compared with traditional orthopedic nursing training, what do you consider to be the advantages and disadvantages of using the AI tool for nursing training? (4) How could the teaching and application of such tools be better integrated into nursing education, and what suggestions do you have?

A researcher trained in qualitative methods conducted all interviews face to face in a quiet conference room. Each interview lasted 30 to 60 minutes. Before each interview began, the researcher explained the study purpose and obtained written informed consent. During the interviews, the researcher used non-leading prompts, observed nonverbal responses, and adjusted follow-up questions according to participants' answers. Interview appointments, including time and place, were arranged with participants in advance.

2.4. Data Collection and Analysis

2.4.1. Quantitative Research

Before training, all participants completed a 13-item self-administered demographic questionnaire that collected age, sex, grade point average, and other baseline characteristics.

The primary outcome was the post-training DDH classification test score, administered only after training to avoid pretest exposure to the image-classification items and potential learning effects. The DDH classification test was developed by the medical team according to the course content and included 12 image-classification items. Each item was scored out of 5 points, including 2 points for

identifying whether DDH was present and 3 points for classifying DDH severity, for a total score of 60. Higher scores indicated greater mastery of DDH classification.

Secondary outcomes included DDH-related nursing assessment, deep learning, self-directed learning, teaching satisfaction, and AI learning intention. The DDH Nursing Assessment Questionnaire was developed by the nursing team and administered before and after training. It included five DDH-related knowledge multiple-choice items and two case-analysis items, with a total score of 40. Two nursing teachers developed the nursing assessment content, and a senior pediatric orthopedic clinician reviewed the disease-related teaching and case materials to support content validity. Higher scores indicated greater mastery of DDH nursing assessment.

Deep learning was assessed using the Scale of Deep Learning for College Students in Chinese Blended Learning Environments [16]. The scale contains 36 items across four dimensions: deep learning motivation, deep learning engagement, deep learning strategies, and deep learning outcomes. Items are scored on a 4-point Likert scale, and total scores range from 36 to 144, with higher scores indicating deeper learning. The scale has demonstrated good reliability and validity, with a *Cronbach's alpha* of 0.971 and acceptable confirmatory factor analysis fit indices.

Self-directed learning was measured using the Self-Directed Learning Instrument, which comprises 20 items across four dimensions: learning motivation, planning and implementation, self-monitoring, and interpersonal communication [17]. Items are scored on a 5-point Likert scale, and total scores range from 20 to 100, with higher scores indicating stronger self-directed learning. The instrument has shown good reliability and content validity (*Cronbach's alpha* = 0.916, content validity index = 0.93).

Teaching satisfaction was assessed after training using a 14-item questionnaire developed from the literature and covering satisfaction and confidence. Items were scored on a 5-point Likert scale, and total scores ranged from 13 to 65, with higher scores indicating greater satisfaction and confidence. The *Cronbach's alpha* coefficient for the teaching satisfaction questionnaire was 0.973.

In the intervention group, willingness to learn AI was assessed using the AI Learning Intention Scale [18]. The scale includes four dimensions and 11 factors: AI basic knowledge, programming self-efficacy, social good, social support, access to learning resources and technology, personal relevance, optimism, digital resilience, behavioral intention to learn AI, actual AI learning behavior, and actual AI system use. Items are scored on a 5-point Likert scale, and higher factor scores indicate stronger AI learning intention. The scale has shown acceptable reliability and validity, with *Cronbach's alpha* coefficients of 0.72 to 0.90 across dimensions, *RMSEA* = 0.04, *CFI* = 0.96, and a content validity index of 0.93.

All questionnaires were distributed, completed, and collected on site by the researchers. Completed questionnaires were checked immediately for completeness. Data were entered into Microsoft Excel 2020 and imported into IBM SPSS Statistics 27.0 for analysis after double-checking.

Categorical variables were summarized as frequencies and percentages and compared using the chi-square test or Fisher's exact test when expected cell counts were small. Normality of continuous variables was assessed using the Shapiro-Wilk or Kolmogorov-Smirnov test. Normally distributed variables were reported as mean $\pm$ standard deviation and compared using independent-samples t tests or repeated-measures analysis of variance, as appropriate. The primary outcome, DDH classification test score, was administered after training and compared between groups using an independent-samples t test. For outcomes measured both before and after training, the time-by-group interaction from repeated-measures analysis was treated as the main test of differential change, followed by simple-effect analyses for descriptive interpretation. Non-normally distributed variables were reported as medians and interquartile ranges. Between-group comparisons used the Mann-Whitney U test, and within-group pre-post comparisons used the Wilcoxon signed-rank test. AILIS factor-level analyses were exploratory; unadjusted p values are reported and should be interpreted cautiously because of multiple comparisons. All analyses were two-sided, and $p < 0.05$ was considered statistically significant. The primary analysis was conducted on participants who completed the assigned intervention and post-training assessments.

2.4.2. Qualitative Research

The interviews were audio-recorded, and nonverbal responses were also documented. One researcher transcribed the recordings within 48 hours and managed using NVivo 15. Another researcher checked the transcripts against the audio recordings for accuracy.

The researcher used Colaizzi's method for analysis: repeated reading, extraction of significant statements, coding of recurring themes, clustering of codes, formulation of an exhaustive description, identification of shared meanings, and participant verification. To ensure rigor, two researchers coded the transcripts independently and resolved disagreements through discussion. Unresolved issues were reviewed by a third researcher. All research team members had received systematic training in qualitative research methods and had qualitative research experience. In addition, preliminary findings were returned to three interviewees, and none raised objections or requested revisions. A nursing education researcher also conducted peer review of the study design, analysis process, and conclusions at the beginning and final stages of the qualitative work.

2.4.3. Mixed-methods integration

Following the separate quantitative and qualitative analyses, the findings were compared to identify meaningful relationships between the quantitative outcomes and qualitative themes. This integration provided a deeper understanding of how AI-assisted image interpretation influenced nursing assessment and decision-making and why certain outcomes improved while others remained unchanged.

3. Results

3.1. Quantitative phase

3.1.1. Participant characteristics

Of the 64 nursing interns assessed for eligibility, four declined participation. Sixty participants were randomized, with 30 allocated to the intervention group and 30 allocated to the control group. Two participants in each group did not complete the assigned training, leaving 56 participants for the final analysis, with 28 participants in each group. The flow of participants through each stage of the randomized trial is detailed in Fig. 1. No harms or unintended adverse events related to the educational intervention were reported in either group. Among the participants included in the analysis, the majority were female ($n = 48, 85.7\%$) and of Han ethnicity ($n = 52, 92.9\%$). The mean age in both groups was approximately 21 years (intervention group, 21.00 ± 0.86 years; control group, 20.93 ± 0.81 years), and the mean duration of clinical practice was also similar (5.57 ± 0.84 months vs. 5.61 ± 0.83 months). Most participants were undergraduate students (intervention group, $n = 21, 75.0\%$; control group, $n = 20, 71.4\%$), and the majority had a GPA between 2.5 and 3.5 (intervention group, $n = 17, 60.7\%$; control group, $n = 18, 64.3\%$). Only a small number of participants had prior experience with pediatric orthopedic rotations (intervention group, $n = 4, 14.3\%$; control group, $n = 5, 17.9\%$). No significant differences were observed between the two groups in demographic characteristics, educational background, family status, income, internship experience, or prior pediatric orthopedic experience (all $p > 0.05$; Table 1).

Table 1
Participant characteristics of nursing interns by group

Characteristic	Intervention group (<i>n</i> = 28)	Control group (<i>n</i> = 28)	Statistic	<i>p</i>
Sex, <i>n</i> (%)			$\chi^2 = 0.000$	1.000 ^a
Male	4 (14.3)	4 (14.3)		
Female	24 (85.7)	24 (85.7)		
Age, years	21.00 ± 0.86	20.93 ± 0.81	<i>t</i> = 0.319	0.751 ^b
Ethnicity, <i>n</i> (%)			$\chi^2 = 0.000$	1.000 ^a
Han Chinese	26 (92.9)	26 (92.9)		
Ethnic minority	2 (7.1)	2 (7.1)		
High school subject track, <i>n</i> (%)			$\chi^2 = 0.287$	0.789 ^a
Science track	14 (50.0)	16 (57.1)		
Humanities track	14 (50.0)	12 (42.9)		
Education level, <i>n</i> (%)			$\chi^2 = 0.091$	1.000 ^a
Bachelor's degree	21 (75.0)	20 (71.4)		
Junior college	7 (25.0)	8 (28.6)		
Residence, <i>n</i> (%)			$\chi^2 = 0.000$	1.000 ^a
Rural	15 (53.6)	15 (53.6)		
Urban	13 (46.4)	13 (46.4)		
Only child, <i>n</i> (%)			$\chi^2 = 0.589$	0.705 ^a
Yes	3 (10.7)	5 (17.9)		
No	25 (89.3)	23 (82.1)		
Father's or primary caregiver's education, <i>n</i> (%)			$\chi^2 = 1.699$	0.686 ^a
Primary school or below	3 (10.7)	5 (17.9)		
Junior high school	12 (42.9)	13 (46.4)		

Characteristic	Intervention group (n = 28)	Control group (n = 28)	Statistic	p
Senior high/technical secondary school	11 (39.3)	7 (25.0)	$\chi^2 = 1.950$	0.618 ^a
Junior college or above	2 (7.1)	3 (10.7)		
Mother's or primary caregiver's education, n (%)				
Primary school or below	6 (21.4)	5 (17.9)		
Junior high school	15 (53.6)	14 (50.0)	$\chi^2 = 1.599$	0.936 ^a
Senior high/technical secondary school	6 (21.4)	5 (17.9)		
Junior college or above	1 (3.6)	4 (14.3)		
Monthly household income, RMB, n (%)				
≤ 8000	10 (35.7)	8 (28.6)	$\chi^2 = 0.287$	0.789 ^a
8001–15000	10 (35.7)	12 (42.9)		
15001–25000	3 (10.7)	3 (10.7)		
25001–40000	0 (0.0)	1 (3.6)		
≥ 40001	0 (0.0)	0 (0.0)	$\chi^2 = 1.183$	0.781 ^a
No stable income	5 (17.9)	4 (14.3)		
Transferred major, n (%)				
Yes	12 (42.9)	14 (50.0)		
No	16 (57.1)	14 (50.0)	$t = -0.160$	0.873 ^b
GPA, n (%)				
≤ 2.5	0 (0.0)	1 (3.6)		
2.5–3.5	17 (60.7)	18 (64.3)		
≥ 3.5	11 (39.3)	9 (32.1)	$\chi^2 = 0.133$	1.000 ^a
Internship duration, months	5.57 ± 0.84	5.61 ± 0.83		
Pediatric orthopedic rotation experience, n (%)				
Yes	4 (14.3)	5 (17.9)		

Characteristic	Intervention group (<i>n</i> = 28)	Control group (<i>n</i> = 28)	Statistic	<i>p</i>
No	24 (85.7)	23 (82.1)		

* a chi-square test or Fisher's exact test, as appropriate for expected cell counts; b independent-samples t test. GPA, grade point average; RMB, Chinese yuan.

3.1.2. Deep learning and self-directed learning

Repeated-measures analysis showed significant time ($F(1,54) = 9.237, p = 0.004$) and between-group effects ($F(1,54) = 4.529, p = 0.038$) for deep learning scores, but the time-by-group interaction was not significant ($F(1,54) = 1.171, p = 0.284$; Table 2). Thus, the analysis did not provide statistical evidence that the pre-post change differed between groups. Simple-effect analysis showed that deep learning scores increased in the intervention group after training (mean change = 7.82, 95% CI [2.31, 13.33], $dz = 0.55, p = 0.005$), whereas the change in the control group was not significant (mean change = 3.72, 95% CI [-1.80, 9.24], $dz = 0.26, p = 0.172$). After the intervention, the intervention group had higher deep learning scores than the control group (MD = 5.64, 95% CI [0.25, 11.03], $d = 0.56, p = 0.040$), but this post-training difference should be interpreted considering the non-significant interaction.

Table 2
Analysis Results for Deep Learning Scale Scores

Analysis	Control group (<i>n</i> = 28)	Intervention group (<i>n</i> = 28)	Statistic	<i>p</i>
Repeated-measures ANOVA				
Time effect			$F = 9.237$	0.004
Between-group effect			$F = 4.529$	0.038
Time × group interaction			$F = 1.171$	0.284
Simple effects				
Pre-intervention score	106.32 ± 7.90	107.86 ± 9.84	$F = 0.415$	0.522
Post-intervention score	110.04 ± 9.38	115.68 ± 10.67	$F = 4.419$	0.040
Within-group change	$F = 1.915; p = 0.172$	$F = 8.492; p = 0.005^a$		

Self-directed learning scores did not differ significantly over time ($p = 0.537$), between groups ($p = 0.096$), or by time-by-group interaction ($p = 0.991$; Table 3). Simple-effect analyses also showed no significant

within-group changes in either the control group ($p = 0.656$) or the intervention group ($p = 0.668$).

Table 3
Analysis results for self-directed learning scores

Analysis	Control group ($n = 28$)	Intervention group ($n = 28$)	Statistic	p
Repeated-measures ANOVA				
Time effect			$F = 0.387$	0.537
Between-group effect			$F = 2.867$	0.096
Time $\times$ group interaction			$F = 0.000$	0.991
Simple effects				
Pre-intervention score	77.14 ± 8.58	79.75 ± 6.94	$F = 1.563$	0.217
Post-intervention score	78.11 ± 8.10	80.68 ± 8.56	$F = 1.333$	0.253
Within-group change	$F = 0.201; p = 0.656$	$F = 0.186; p = 0.668^a$		

* a Within-group comparison with the pre-intervention score. The between-group effect was $p = 0.096$.

3.1.3. AI learning intention in the intervention group

In the intervention group, several AI learning intention factors showed within-group increases after training: Programming self-efficacy, social good, digital resilience, behavioral intention to learn AI, and actual AI learning behavior (all $p < 0.05$; Table 4). Because 11 AILIS factors were examined in an exploratory within-group analysis without multiplicity adjustment and without a control-group comparison, these findings should be interpreted as hypothesis-generating. Scores for AI basic knowledge, social support, access to learning resources and technology, personal relevance, optimism, and actual AI system use did not change significantly (all $p > 0.05$).

Table 4
AI Learning Intention Scale scores before and after intervention in the intervention group

AILIS factor	Pre-intervention	Post-intervention	<i>t</i>	<i>p</i>
AICK	16.25 ± 2.01	16.46 ± 0.92	-0.381	0.706
ProgE	10.14 ± 3.79	12.57 ± 3.87	-2.775	0.010
SG	14.25 ± 2.19	15.64 ± 1.68	-2.731	0.011
SN	11.04 ± 1.43	11.71 ± 1.41	-1.846	0.076
Aces	14.64 ± 2.45	15.61 ± 1.66	-1.533	0.137
Rel	15.86 ± 1.63	16.36 ± 1.75	-1.064	0.297
Op	15.18 ± 2.18	16.04 ± 1.80	-1.441	0.161
Res	15.46 ± 2.12	16.25 ± 1.80	-2.274	0.031
BI	15.18 ± 2.51	16.46 ± 1.86	-2.253	0.033
AL	10.21 ± 2.38	11.57 ± 1.87	-2.253	0.033
AU	12.43 ± 1.53	12.57 ± 1.32	-0.335	0.740

* AILIS, Artificial Intelligence Learning Intention Scale; AICK, AI basic knowledge; ProgE, programming self-efficacy; SG, social good; SN, social support; Aces, access to learning resources and technology; Rel, personal relevance; Op, optimism; Res, digital resilience; BI, behavioral intention to learn AI; AL, actual AI learning behavior; AU, actual AI system use. P values are unadjusted exploratory within-group comparisons.

3.1.4. DDH-related nursing assessment, classification, and satisfaction

For DDH-related nursing test scores, repeated-measures analysis showed a significant between-group effect, but the time effect and time-by-group interaction were not significant (Table 5). Therefore, the analysis did not demonstrate a statistically significant difference in pre-post change between groups. Simple-effect analysis showed no significant pre-post change in the control group ($p = 0.538$), whereas the intervention group showed a borderline increase after training ($p = 0.055$). After the intervention, nursing test scores were higher in the intervention group than in the control group ($MD = 3.72$, 95% CI [0.96, 6.48], $d = 0.72$, $p = 0.009$), but this post-training difference should be interpreted cautiously because the interaction effect was not significant.

After training, the intervention group also scored higher than the control group on the DDH classification test ($MD = 4.07$, 95% CI [0.17, 7.97], $d = 0.56$, $p = 0.041$; Table 5). Teaching satisfaction did not differ significantly between groups ($p = 0.552$; Table 6).

Table 5
DDH-related nursing test and classification test scores

Outcome/analysis	Control group (<i>n</i> = 28)	Intervention group (<i>n</i> = 28)	Statistic	<i>p</i>
Nursing test: repeated-measures ANOVA				
Time effect			<i>F</i> = 3.333	0.073
Between-group effect			<i>F</i> = 6.076	0.017
Time × group interaction			<i>F</i> = 0.901	0.347
Nursing test: simple effects				
Pre-intervention score	18.86 ± 5.85	20.71 ± 6.16	<i>F</i> = 1.338	0.253
Post-intervention score	19.71 ± 4.86	23.43 ± 5.41	<i>F</i> = 7.301	0.009
Within-group change	<i>F</i> = 0.384; <i>p</i> = 0.538	<i>F</i> = 3.850; <i>p</i> = 0.055a		
DDH classification test after intervention	34.88 ± 7.11	38.95 ± 7.44	<i>t</i> = 2.093	0.041

* a Within-group comparison with pre-intervention score. The DDH classification test was administered after training and compared between groups using an independent-samples *t* test.

Table 6
Teaching satisfaction scores after intervention

Outcome	Intervention group (<i>n</i> = 28)	Control group (<i>n</i> = 28)	Statistic	<i>p</i>
Teaching satisfaction	52.00 (52.00, 57.75)	53.00 (51.25, 64.50)	<i>Z</i> = -0.594	0.552

3.2. Qualitative phase

Ten nursing interns from the intervention group participated in face-to-face interviews. Participants were aged 20–22 years, and two had previous pediatric orthopedic rotation experience. To protect participant anonymity, demographic characteristics are reported only in aggregate form (Table 7). The interviews lasted 30 to 60 min, with a total interview duration of 376 min and approximately 45,000 Chinese characters transcribed.

Table 7
Aggregate characteristics of interview participants (*n* = 10)

Characteristic	Summary
Age range, years	20–22
Sex, <i>n</i> (%)	
Female	8 (80.0)
Male	2 (20.0)
Education level, <i>n</i> (%)	
Bachelor's degree	8 (80.0)
Junior college	2 (20.0)
*GPA, <i>n</i> (%)	
2.5–3.5	5 (50.0)
≥3.5	5 (50.0)
Previous pediatric orthopedic rotation, <i>n</i> (%)	
Yes	2 (20.0)
No	8 (80.0)
DDH classification test score, mean ± SD	41.40 ± 8.75
DDH classification test score, range	30.0–54.0

* GPA, grade point average.

Four themes and ten subthemes were identified from the interview data (Table 8). Participants described that AI-assisted training supported knowledge integration, clinical reasoning, professional confidence, and professional identity. They also reported greater willingness to learn both DDH-related knowledge and AI-related skills. Practical advantages of the AI tool included efficiency, intuitive visualization, interactivity, immediate feedback, and accessibility. Reported limitations included narrow classification content, potential dependence on tool outputs, weakened critical thinking, and limited support for advanced learning. Participants further highlighted implementation challenges, including the tension between autonomous learning and high learning demands, the need for continuous alignment between educational content and evolving technology, and the need for more diverse learning formats.

Table 8
Themes and subthemes from qualitative interviews

Theme	Subtheme	Description	Representative quotation
AI-assisted training supported nursing interns	Knowledge integration and clinical reasoning	Imaging-focused training helped students link DDH classification, disease severity, surgical decisions, and nursing care.	"Strengthening imaging learning allowed us to see disease severity and classification, understand why surgery was performed, and support later nursing care." (S3)
AI-assisted training supported nursing interns	Professional confidence	The tool helped students judge patient conditions and develop more targeted care plans and health education.	"If the tool helps me confirm the patient's condition, I can distinguish treatment options and nursing measures more clearly." (S9)
AI-assisted training supported nursing interns	Professional identity	Students felt more able to participate in clinical decision processes rather than only execute care tasks.	"I am not just an executor; I can work with doctors as a collaborator in the care process." (S4)
Intelligent training changed learning willingness	Willingness to learn DDH-related knowledge	The novelty of imaging and AI-assisted learning increased interest and reduced resistance to completing learning tasks.	"With this tool, I would be more active and more willing to complete the assignments and tests." (S9)
Intelligent training changed learning willingness	Willingness to transfer AI-related skills	Exposure to the tool increased awareness of intelligent clinical tools, although some students remained cautious about accuracy.	"It promoted students' understanding of intelligent tools and new methods used in clinical practice." (S10)
Practical advantages and potential limitations of the AI tool	Tool advantages	Students emphasized efficiency, accuracy, convenience, intuitive visualization, interaction, immediate feedback, and practical value for beginners.	"The tool is intelligent, analyzes accurately, and identifies the image clearly." (S2)
Practical advantages and potential limitations of the AI tool	Tool limitations	Students noted limited classification methods, possible dependence on tool outputs, weakened critical thinking, and limited usefulness for advanced learning.	"After having this tool, I feel I may rely on it, and my own active thinking may decrease." (S9)
Implementation challenges in	Autonomy-demand	Autonomous learning was difficult for some students	"The course requires self-directed completion, but
<i>Note: Representative quotations were translated from the original Chinese interview transcripts.</i>			

Theme	Subtheme	Description	Representative quotation
training design	tension	because of busy internship schedules and weak anatomy foundations.	internships are busy, which may hinder the course." (S5)
Implementation challenges in training design	Dynamic integration of technology and education	Students noted that both AI tools and clinical teachers need continuous updating as medical knowledge evolves.	"The tool must be continuously improved, and we need to keep up with the latest standards." (S1)
Implementation challenges in training design	Need for diverse learning formats	Students wanted real-case observation, offline practice, and structured learning materials such as mind maps.	"A mind map at the beginning would help us build a systematic framework before learning details." (S7)
<i>Note: Representative quotations were translated from the original Chinese interview transcripts.</i>			

4. Discussion

4.1. AI-assisted education bridges imaging judgment and nursing decision-making

This mixed-methods randomized controlled trial suggested that incorporating an AI-based DDH diagnostic tool into orthopedic nursing education was associated with higher post-training scores in DDH classification, DDH-related nursing assessment, and deep learning. However, for the secondary outcomes measured before and after training, the time-by-group interactions were not statistically significant. Therefore, these findings should be interpreted as preliminary evidence of post-training group differences rather than definitive evidence of greater improvement over time. The qualitative interviews suggested a plausible educational mechanism: AI-assisted tools may support nursing students' test performance by strengthening knowledge integration and clinical reasoning. With the support of this tool, nursing interns were able to interpret radiographic images, identify DDH classification, disease severity, and potential treatment options, and connect these findings with priorities for nursing assessment, nursing decisions, and care planning. This process may help students move from fragmented image recognition toward more integrated clinical reasoning. One long-standing limitation of traditional teaching methods is limited support for knowledge integration [19]. The use of AI in education may help address this limitation. However, because self-directed learning is a behavioral skill that relies on time management and sustained effort, and because the interns had busy schedules and limited backgrounds in anatomy, self-directed learning may not improve substantially in the short term.

The observed improvements may be explained by the capacity of AI-assisted instruction to support knowledge integration, reduce cognitive burden, and provide timely feedback. A study reported that generative AI can facilitate case-based learning through structured reasoning and real-time feedback

[20]. The present findings suggest that AI tools in imaging education can transform images from isolated visual materials into structured cues for clinical reasoning by visualizing anatomical landmarks and classification results. This may help nursing interns integrate fragmented imaging knowledge into nursing pathways. Furthermore, by reducing the external cognitive load associated with interpreting unstructured data, AI may enable learners to devote more cognitive resources to diagnostic reasoning, image-pathology correlation, and the translation of imaging findings into clinical decisions [21]. AI can also provide instant feedback on interpretation results, enabling students to receive timely responses for reflection and facilitating the application of imaging findings [22].

4.2. AI-assisted education strengthens AI readiness and learning engagement

The intervention group showed selective nominal increases in several AI learning intention factors. After training, interns in the intervention group had higher programming self-efficacy, social-good orientation, digital resilience, behavioral intention to learn AI, and actual AI learning behavior scores on unadjusted within-group tests, whereas AI basic knowledge, social support, access to learning resources and technology, personal relevance, optimism, and actual AI system use remained unchanged. Because these factor-level analyses were exploratory and were not adjusted for multiple comparisons, this pattern should be viewed as a signal for future confirmatory studies rather than conclusive evidence of broad AI readiness improvement. The results suggest that a short, task-embedded AI intervention may be more likely to influence learners' confidence, adaptive readiness, and willingness to engage with AI than broader contextual perceptions or foundational AI knowledge [23, 24].

The qualitative findings help explain this selective effect. The interviewed interns believed that the AI tool offered advantages such as efficiency, intuitive visualization, interactivity, and timely feedback. This image-interpretation-based learning experience may have helped students recognize the practical value of AI in solving clinically situated problems, rather than viewing AI merely as an abstract technical concept [25]. As a result, students may be more likely to develop self-efficacy, the belief that they can access and learn to use AI tools. This may enhance their adaptability and willingness to explore intelligent technologies [26]. At the same time, by applying AI tools to disease assessment and care decision-making, interns may strengthen their commitment to public service and recognize that learning about AI has professional value in improving clinical care. This recognition may enhance their professional confidence and professional identity [27]. No significant changes were observed in students' foundational knowledge of AI or their practical use of AI systems. This finding indicates that short-term exposure to a single diagnostic tool is insufficient to help students establish a systematic framework of AI knowledge and is unlikely to translate into consistent AI usage behaviors [28].

4.3. Risks and pedagogical implications of AI-assisted education

However, AI-assisted orthopedic education also carries certain risks. First, users may become overly reliant on the results generated by the tool. Participants noted that access to automated classification

results may reduce the motivation to conduct independent visual analysis. Previous research supports this concern, showing that while learners can benefit from AI guidance, some may also develop automation bias, a tendency to over-rely on algorithmic outputs and reduce independent verification. This may undermine independent critical thinking. Second, when AI is applied to teaching, it may implicitly assume that learners possess a solid foundation and comparable learning abilities [29]. Learners with weaker foundations may be unable to interpret AI feedback effectively or transfer and deepen their learning based on AI results. As a result, they may benefit far less from the tool than their peers in more favorable circumstances [30]. This interpretation is consistent with the present finding that self-directed learning did not improve significantly among the interns.

These risks offer important insights for the future of AI-assisted education. First, in orthopedic nursing imaging education, learners should conduct a preliminary independent assessment, clarify the anatomical basis and nursing implications of their judgments, and then compare their reasoning with AI feedback. This process maintains independent observation and critical verification while allowing AI to provide timely corrections and reflective feedback [6]. Second, AI-assisted instruction should shift from a one-size-fits-all model to a tiered instructional design. Educators should assess learners' foundational imaging knowledge, self-directed learning abilities, and AI literacy before training and then provide differentiated learning plans [31]. Students with weak foundations or limited learning abilities may require systematic guidance and instructor support in interpreting AI feedback, whereas more capable learners can engage in open-ended case studies and critically evaluate algorithmic limitations.

4.4. Limitations

This study has several limitations. First, it was conducted at a single tertiary hospital using a relatively small convenience sample of nursing interns, which may limit the generalizability of the findings and reduce the statistical power to detect group-by-time interaction effects. Second, because the final analysis included participants with complete post-training data rather than the full randomized sample, potential attrition bias cannot be fully excluded. In addition, the DDH classification test was administered only after training to avoid pretest exposure and potential learning effects, so baseline-adjusted analysis was not possible for this outcome. Third, the intervention was of relatively short duration, focused primarily on DDH, and incorporated only one AI-assisted diagnostic tool. Future multicenter studies with larger samples should broaden the instructional content and clinical application contexts and include longer follow-up periods to evaluate the sustainability of learning effects and their transferability to clinical practice.

5. Conclusion

This mixed-methods trial suggests that integrating an AI-based DDH diagnostic tool into orthopedic nursing education was associated with higher post-training DDH classification, nursing assessment, and deep learning scores. Qualitative findings suggest that AI-assisted education may help nursing interns connect anatomical landmarks, disease severity, treatment plans, and nursing decisions, thereby supporting their transition from fragmented image recognition to clinically oriented reasoning. The

intervention group also showed nominal increases in several AI learning intention factors, including programming self-efficacy, digital resilience, social-benefit orientation, behavioral intention to learn AI, and actual AI learning behavior.

However, AI-assisted education may also have limitations, such as potentially affecting learners' independent thinking and exacerbating disparities among learners. Future applications should integrate AI feedback with learners' initial reasoning, instructor supervision, and tiered instructional design to reduce overreliance, support learners with varying levels of prior knowledge, and promote the responsible integration of AI into orthopedic nursing education.

Abbreviations

(1) AI

Artificial intelligence

(2) DDH

Developmental dysplasia of the hip

Declarations

Ethics approval and consent to participate:

This study was approved by the Ethics Committee of the Third Affiliated Hospital of Southern Medical University (2024-ER-112). All participants signed paper informed consent forms. All procedures involving human participants were performed in accordance with the ethical standards of the relevant institutional and/or national research committee and with the Declaration of Helsinki and its subsequent amendments or comparable ethical standards.

Consent for publication:

Not applicable.

Competing interests:

The authors have no relevant financial or non-financial interests to disclose.

Funding:

This work was supported by the 2025 Education and Teaching Research Project of the Third Affiliated Hospital of Southern Medical University (grant number: JXY202517).

Author Contribution

Z.X.Y. conducted the investigation and drafted the original manuscript. G.L. conceptualized the study, developed the methodology, curated the data, and performed the formal analysis. Y.Y.L. contributed to the investigation. S.T.H. and Y.K.C. contributed to the formal analysis. J.J.P., G.Y.D., J.H.Y., and Y.C. provided resources. H.O. and Y.T.L. contributed to data curation. F.Q. reviewed and edited the manuscript. J.H.Y. administered the project, and Y.C. supervised the study. All authors read and approved the final manuscript.

Acknowledgements:

λ The authors would like to thank all the interns who participated in this study.

Data Availability

The datasets generated and analyzed during the current study are not publicly available due to participant privacy concerns but are available from the corresponding author on reasonable request.

References

1. Cui J, Liang Y, Wang Y, Guo F, Yang D, Liao Y. Impact of precision nursing intervention based on DCO model on joint function and quality of life of elderly hip arthroplasty patients. *J Orthop Surg Res.* 2025;20:421–30. <https://doi.org/10.1186/s13018-025-05820-1>.
2. Neep MJ, Steffens T, Eastgate P, McPhail SM. Evaluating the effectiveness of intensive versus non-intensive image interpretation education for radiographers: A randomised control trial study protocol. *J Med Radiat Sci.* 2018;65:63–70. <https://doi.org/10.1002/jmrs.264>.
3. Yoshioka-Maeda K, Matsumoto H, Honda C, Kinjo T, Aoki K, Hattori T, et al. Integrating web-based and objective structured clinical examination training programs for nurses on ultrasound hip screening during neonatal and infant home visits: A case study. *BMC Med Educ.* 2025;25:1007. <https://doi.org/10.1186/s12909-025-07519-7>.
4. Kurul R, Inal B, Diramali M, Ozer H, Erdogan KE. Evaluating the educational impact of tomography-based high-resolution 3D-printed distal radius fracture models. *BMC Med Educ.* 2025;25:1706. <https://doi.org/10.1186/s12909-025-08164-w>.
5. Mohd Salim NA, Azman A, Sulaiman PS, Roslan NS, Hod R, Adam SK. Development and validation of virtual patients for assessing history taking and communication skills among medical students. *BMC Med Educ.* 2026;26:264–73. <https://doi.org/10.1186/s12909-025-08503-x>.
6. Tee Y-S, Liao C-A, Kuo L-W, Hsu C-P, Wang C-C, Fu C-Y, et al. Artificial intelligence-assisted training for rib fracture interpretation: A prospective study in undergraduate medical students. *World j emerg surg: WJES.* 2026;21:14. <https://doi.org/10.1186/s13017-026-00678-y>.

7. Lin Y, Luo Z, Ye Z, Zhong N, Zhao L, Zhang L, et al. Applications, challenges, and prospects of generative artificial intelligence empowering medical education: Scoping review. *JMIR Med Educ*. 2025;11:e71125–71125. <https://doi.org/10.2196/71125>.
8. Gordon M, Daniel M, Ajiboye A, Uraiby H, Xu NY, Bartlett R, et al. A scoping review of artificial intelligence in medical education: BEME guide 84. *Med Teach*. 2024;46:446–70. <https://doi.org/10.1080/0142159X.2024.2314198>.
9. Zeng J, Qi W, Shen S, Liu X, Li S, Wang B, et al. Embracing the future of medical education with large language model–based virtual patients: Scoping review. *J Med Internet Res*. 2025;27:e79091. <https://doi.org/10.2196/79091>.
10. Chan KS, Zary N. Applications and challenges of implementing artificial intelligence in medical education: integrative review. *JMIR Med Educ*. 2019;5:e13930. <https://doi.org/10.2196/13930>.
11. Najjar R. Redefining radiology: A review of artificial intelligence integration in medical imaging. *Diagnostics*. 2023;13:2760. <https://doi.org/10.3390/diagnostics13172760>.
12. Faul F, Erdfelder E, Lang A-G, Buchner A. G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39:175–91. <https://doi.org/10.3758/bf03193146>.
13. Sia C-H, Ng S, Hoon D, Soong J, Ignacio J, Kowitlawakul Y. The effectiveness of collaborative teaching in an introductory online radiology session for master of nursing students. *Nurse Educ Today*. 2021;105:105033. <https://doi.org/10.1016/j.nedt.2021.105033>.
14. Chen J, Fan Xi, Zhen C, Peng Y, Liang C, C S, et al. Enhancing YOLO5 for the assessment of irregular pelvic radiographs with multimodal information. *J imaging Inf Med*. 2024. <https://doi.org/10.1007/s10278-024-00986-2>. 37.
15. Li G, Li X, Li P, Wei Q, Lijuan LI, Lei, Wang, et al. A Preliminary Exploration of Online Education Based on Ebbinghaus’s Theory of Memory: A Case Study of Speech-Language Pathology Instruction. *Chin J Rehabil Med*. 2022;1234–7. <https://doi.org/10.3969/j.issn.1001-1242.2022.09.014>.
16. Li Y, Su D, Li Q, Ren Y. Development of a Deep Learning Scale for College Students in Hybrid Learning Environments. *E-educ Res*. 2018;39:94–101. <https://doi.org/10.13811/j.cnki.eer.2018.12.013>.
17. Cheng S-F, Lee-Hsieh J, Turton MA, Lin K-C. Validation of self-directed learning instrument and establishment of normative data for nursing students in taiwan: Using polytomous item response theory. *J Nurs Res*. 2014;22:90. <https://doi.org/10.1097/JNR.0000000000000027>.
18. Chai C, Yu D, King RB, Zhou Y. Development and validation of the artificial intelligence learning intention scale (AILIS) for university students. *Sage Open*. 2024;14:21582440241242188. <https://doi.org/10.1177/21582440241242188>.
19. Yang Z, Wen J, Bi X, Lin J, Zhao Q, Zhang P, et al. Development and application of an artificial intelligence agent-based case teaching model for health assessment: A quasi-experimental study. *Nurse Educ Today*. 2026;164:107142. <https://doi.org/10.1016/j.nedt.2026.107142>.
20. Abidi SH, Almazan J, Fabiyi O, Zehra F, Tariq M. AI-supported case-based learning in medical education: A comprehensive scoping review. *Front Med*. 2026;13.

<https://doi.org/10.3389/fmed.2026.1798097>.

21. Li J, Zhao H. Workflow-embedded AI as a cognitive scaffold: A randomized trial on knowledge retention and diagnostic competency in undergraduate radiology education. *Eur J Radiol Open*. 2026;16. <https://doi.org/10.1016/j.ejro.2026.100724>.
22. Wentzell S, Moran L, Dobranowski J, Levinson A, Hannigan A, Dunne CP, et al. E-learning for chest x-ray interpretation improves medical student skills and confidence levels. *BMC Med Educ*. 2018;18:256. <https://doi.org/10.1186/s12909-018-1364-2>.
23. Kwak Y, Ahn J-W, Seo YH. Influence of AI ethics awareness, attitude, anxiety, and self-efficacy on nursing students' behavioral intentions. *BMC Nurs*. 2022;21:267. <https://doi.org/10.1186/s12912-022-01048-0>.
24. Li X, Jiang MY, Jong MS, Zhang X, Chai C. Understanding medical students' perceptions of and behavioral intentions toward learning artificial intelligence: A survey study. *Int J Environ Res Public Health*. 2022;19:8733. <https://doi.org/10.3390/ijerph19148733>.
25. Cheng C-T, Chen C-C, Fu C-Y, Chaou C-H, Wu Y-T, Hsu C-P, et al. Artificial intelligence-based education assists medical students' interpretation of hip fracture. *Insights into Imaging*. 2020;11:119. <https://doi.org/10.1186/s13244-020-00932-0>.
26. Al Omari O, Alshammari M, Al Jabri W, Al Yahyaei A, Aljohani KA, Sanad HM, et al. Demographic factors, knowledge, attitude and perception and their association with nursing students' intention to use artificial intelligence (AI): A multicentre survey across 10 arab countries. *BMC Med Educ*. 2024;24:1456. <https://doi.org/10.1186/s12909-024-06452-5>.
27. Bodur G, Cakir H, Turan S, Seren AKH, Goktas P. Artificial intelligence in nursing practice: A qualitative study of nurses' perspectives on opportunities, challenges, and ethical implications. *BMC Nurs*. 2025;24:1263. <https://doi.org/10.1186/s12912-025-03775-6>.
28. Rincón EHH, Jimenez D, Aguilar LAC, Flórez JMP, Tapia ÁER, Peñuela CLJ. Mapping the use of artificial intelligence in medical education: A scoping review. *BMC Med Educ*. 2025;25:526. <https://doi.org/10.1186/s12909-025-07089-8>.
29. Mendoza NB, Xiong Y, Yan Z. Making sense of AI feedback: How students' feedback literacy moderates the link between ChatGPT acceptance and self-regulated learning. *Educ Psychol*. 2026;1–23. <https://doi.org/10.1080/01443410.2026.2641521>.
30. Turney J, Young TM, Chauhan DR, Beeharry R, Mahmud M. AI use for medical students: Impact on clinical skill acquisition and retention. A systematic review. *Adv Med Educ Pract*. 2026;17:1–13. <https://doi.org/10.2147/AMEP.S583763>.
31. Duong MT, Rauschecker AM, Rudie JD, Chen P-H, Cook TS, Bryan RN, et al. Artificial intelligence for precision education in radiology. *Br J Radiol*. 2019;92:20190389. <https://doi.org/10.1259/bjr.20190389>.

Figures

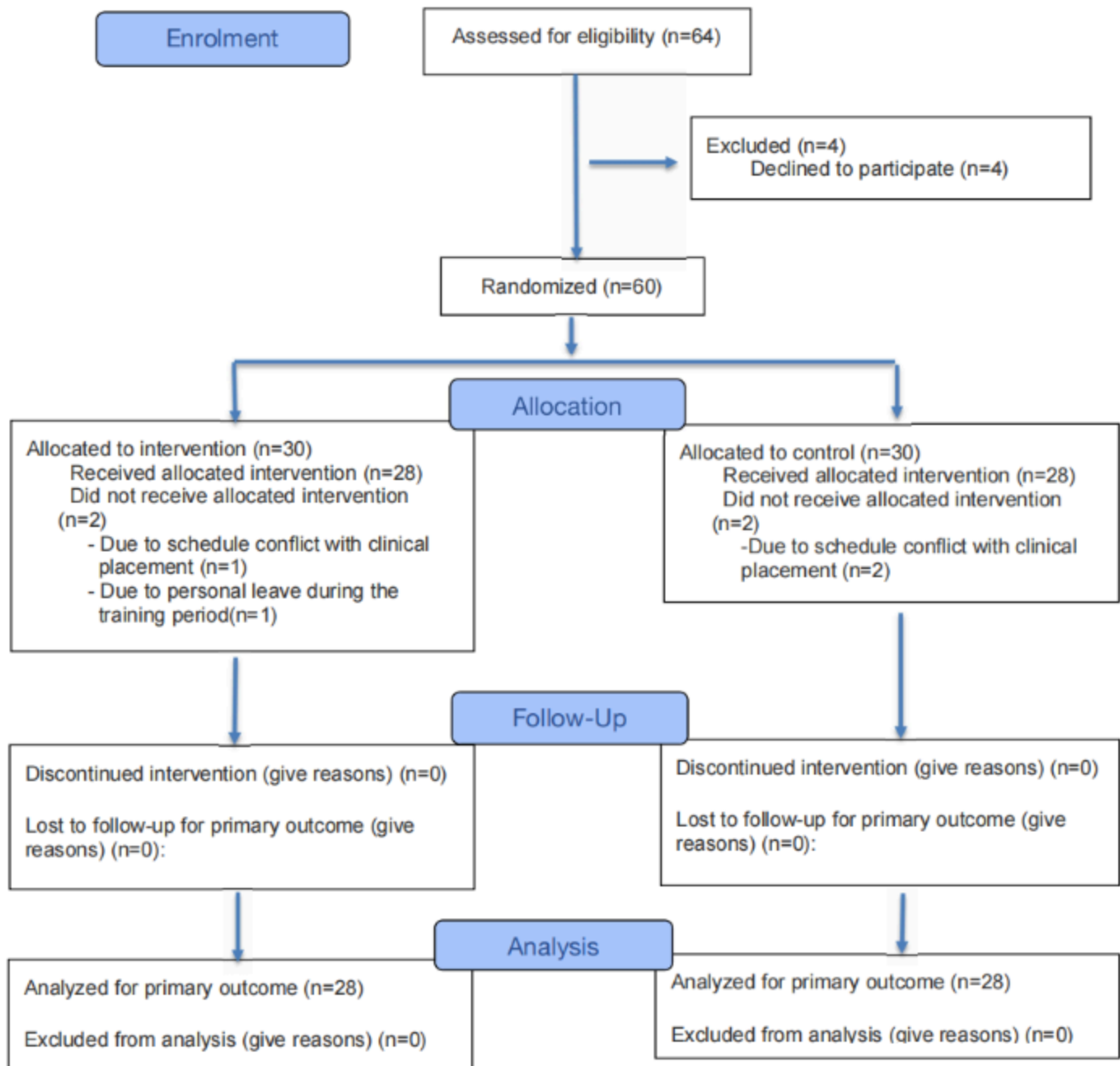


Figure 1

CONSORT Flow diagram

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFileEnglishlanguageversionsofthestudydevelopedinstruments.docx](#)