Fung et al. BMC Nursing (2025) 24:934
https://doi.org/10.1186/s12912-025-03492-0

BMC Nursing

RESEARCH

Open Access

Effects of generative artificial intelligence (GenAI) patient simulation on perceived clinical competency among global nursing undergraduates: a cross-over randomised controlled trial

Tai Chun John Fung[1*], Siu Ling Chan[1], Choi Fung Mabel Lam[1], Chung Yan Lam[1], Christopher Chi Wai Cheng[3], Man Hin Lai[1], Cheuk Chun Joseph Ho[1], Siu Lun Au[1], Lok Yi Mak[4], Sophia Hu[5], Supapak Phetrasuwan[6], Jumpee Granger[7], Jung Min Yoon[8], Gulzar Malik[9], Clara Cabrera Moreno[10], Man Hei Patrick Kwok[1,2] and Chia-Chin Lin[1*]

Abstract

Background This study compared scenario-based generative artificial intelligence (GenAI) patient simulation with immersive 360° virtual reality (VR) simulation in terms of perceived clinical competence, cultural awareness, AI readiness, and simulation effectiveness among nursing students.

Methods This cross-over randomised controlled study design was conducted from June 2024 to August 2024. Forty-four undergraduate nursing students from years 1–3 were randomised to receive either GenAI patient simulation (Group B) or 360° VR simulation (Group A) with a one-week washout period. Five self-reported questionnaires were used to measure clinical competency: the Clinical Competence Questionnaire (CCQ), Cultural Awareness Scale (CAS), Medical Artificial Intelligence Readiness Scale for Medical Students (MAIRS-MS), Simulation Effectiveness Tool – Modified Questionnaire (SET-M), and a demographic questionnaire.

Results Both interventions significantly improved clinical competence, cultural awareness, and AI readiness. When administered first, GenAI patient simulation demonstrated greater initial effects on clinical competence and AI readiness compared to the 360° VR simulation, though both groups achieved similar improvements by study completion. At T1, Group B (receiving GenAI) demonstrated significantly larger improvements in CCQ total score [47.68 (95% CI: 36.68, 58.68), $p < 0.001$] compared to Group A (receiving 360° VR) [24.95 (95% CI: 13.96, 35.95), $p < 0.001$], with significant between-group difference [16.59 (95% CI: 2.77, 30.41), $p = 0.020$]. At T2 (post-crossover), both groups maintained significant improvements. For MAIRS-MS (measured at baseline and following each group's GenAI exposure), Group B showed improvement from baseline to T1 [30.18 (95% CI: 23.35, 37.01), $p < 0.001$] while

*Correspondence:
Tai Chun John Fung
bigjohn@hku.hk
Chia-Chin Lin
cclin@hku.hk
Full list of author information is available at the end of the article

Group A showed improvement from baseline to T2 [16.64 (95% CI: 9.80, 23.47), $p < 0.001$], with significant between-group difference [12.09 (95% CI: 4.43, 19.75), $p = 0.003$]. Both groups experienced changes in CAS scores, though between-group differences were not statistically significant. For SET-M, most participants (75%) felt debriefing contributed to their learning, and 68.2% reported increased confidence in nursing assessment skills.

Conclusions The findings provide preliminary evidence of its effectiveness in enhancing perceived clinical outcomes among nursing students. Both 360° VR simulation and GenAI patient simulation may serve as effective teaching tools; however, GenAI patient simulation appeared to demonstrate a greater initial effect on clinical competence and AI readiness, although both interventions proved effective across all measured domains.

Clinical trial registration/number Not applicable.

Keywords Generative artificial intelligence, Patient simulation, Clinical reasoning, 360-degree virtual reality, Clinical competence, Medical language, Nursing education, Randomised controlled trial

Introduction

The integration of artificial intelligence (AI) into healthcare has prompted a growing emphasis in the literature on the necessity of AI readiness among nursing students and professionals. AI readiness encompasses the knowledge, skills, and attitudes required for the practical application of AI in healthcare, including preventive measures, diagnostics, treatment, and comprehensive care delivery [1]. However, the potential for AI to perpetuate biases and adversely affect patient care has been noted, emphasizing the need for rigorous training and oversight [2]. To mitigate these risks, nursing curricula should proactively foster AI readiness through targeted education on AI fundamentals, its strengths and limitations, and ethical considerations. The literature supports curricular innovations, such as dedicated coursework and simulation-based learning with GenAI, to build foundational AI readiness and reduce biases in healthcare [3, 4]. Further research is needed to systematically evaluate various pedagogical approaches utilizing AI and their long-term effects.

Background

The increasing integration of artificial intelligence (AI) into healthcare has underscored the necessity for nursing students and professionals to develop AI readiness [5]. This readiness is crucial, as AI technologies, including virtual reality (VR) and conversational agents, are becoming increasingly integral to nursing education, thereby enhancing learning experiences by providing immersive and interactive environments [6]. Notably, 360° VR videos simulate health and social care scenarios that are typically inaccessible in traditional educational settings, thereby significantly improving communication skills and self-efficacy in AI-enabled simulations. However, concerns exist regarding AI's perceived human-like qualities, which could impact the authenticity of these simulations [7, 8].

Building on this foundation, pilot studies have demonstrated that integrating 360° VR simulation with structured debriefing through learning management systems (LMS) can enhance students' clinical competence and confidence in decision-making [9, 10]. This approach not only fosters a deeper understanding of clinical scenarios but also prepares students for the complexities of AI-driven healthcare environments. The emergence of Generative AI (GenAI) patient simulation represents a novel trend in nursing education, necessitating targeted education on AI fundamentals, its strengths, limitations, and ethical considerations [11].

The literature further emphasizes the importance of stakeholder engagement and a comprehensive understanding of AI's clinical role [12]. This includes integrating coursework that addresses AI biases, ensuring that future nurses are equipped to navigate the ethical and practical challenges posed by AI technologies [13]. Nursing practice requires providing care to patients from diverse cultural backgrounds. However, there are concerns that unintentional biases inherent in AI systems may influence learners' cultural awareness when caring for clients [14]. While text-to-text GenAI tools like ChatGPT offer educational benefits, caution is advised due to potential risks, including AI hallucination, misinformation, and cultural sensitivity issues [15]. These findings advocate for urgent curricular reform to prepare nurses for AI practice, highlighting the need for a balanced approach that leverages AI's potential while mitigating its risks [16, 17].

Simulation-based GenAI learning has been shown to enhance AI readiness and reinforce core nursing concepts [18]. Students exposed to GenAI-generated cases exhibit improved clinical reasoning compared to those trained through traditional methods [19]. This suggests that GenAI can play a pivotal role in preparing nursing students for real-world clinical challenges. However, further research is imperative to optimize the integration of GenAI into nursing education while ensuring that training remains bias-free.

In summary, the adoption of AI in nursing education is accelerating, necessitating improved competency among students and faculty. Current literature supports curricular innovations such as dedicated coursework and GenAI simulations to build AI readiness and mitigate biases. Future research should systematically evaluate various AI pedagogical approaches and their long-term impacts, ensuring that nursing education evolves in tandem with technological advancements in healthcare.

Research question

Does participation in generative artificial intelligence (GenAI) patient simulation, compared to immersive 360° virtual reality (VR) simulation, lead to greater improvements in clinical competence, cultural awareness, and AI readiness among global undergraduate nursing students?

Methods

Study design

This was a cross-over randomised controlled study with two intervention arms: GenAI and 360° VR simulation. The participants were allocated equally to the study arms.

Study participants

Forty-four undergraduate nursing students from Hong Kong, Taiwan, Thailand, Spain, South Korea, and Australia, in their first, second, and third years of study, participated. Subgroups were created, each consisting of three undergraduate nursing students from various academic years. A total of 22 participants were randomly assigned to Group A (who received the 360° VR simulation first and GenAI second), and the other 22 participants were assigned to Group B (who received GenAI first and the 360° VR simulation second). The allocation process was randomised with a 1:1 allocation ratio. This sequence was generated by an independent researcher to ensure unbiased assignment and was concealed from both the participants in the two groups and the researchers involved in the study. Moreover, all participants provided written informed consent before the study, and the voluntary nature of participation and the confidentiality of the data were strongly emphasized throughout the process. This study was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (IRB number: UW 24–396).

Data collection

CCQ

The 47-item CCQ, originally developed by Liou, was utilized in this study to assess participants' self-perceived clinical competency at both T0 and T1 [20]. The questionnaire consists of four competency components, namely nursing professional behaviours (NP; 16 items, score range: 16–90), general performance (GP; 13 items, score range: 13–65), core nursing skills (CNS; 12 items, score range: 12–60), and advanced nursing skills (ANS; 6 items, score range: 6–30). The items are scored using a 5-point Likert scale (1 = 'Do not have a clue'; 2 = 'Know in theory, but not confident at all in practice'; 3 = 'Know in theory, can perform some parts in practice independently, and need supervision to be readily available'; 4 = 'Know in theory, competent in practice, need contactable sources of supervision'; 5 = 'Know in theory, competent in practice without supervision'). The total score ranges from 47 to 235 [20]. A higher score indicates higher perceived competence. Cronbach's alpha for the CCQ was 0.98, with the reliability of each component ranging from 0.87 to 0.95 [21].

CAS

The CAS was designed by Rew to measure the cultural awareness of nursing students [22]. The CAS is based on the pathway model and comprises 36 items, each rated on a 7-point Likert scale (1 = 'strongly disagree' to 7 = 'strongly agree') and divided into five subscales: general education experience (GEE; 13 items, score range: 13–91), cognitive awareness (CA; 9 items, score range: 9–63), research issues (RI; 4 items, score range: 4–28), behavior or comfort with interactions (BOCWI; 6 items, score range: 6–42), and patient care or clinical issues (PCOCI; 4 items, score range: 4–28). The total score ranges from 36 to 252. A higher score on each subscale indicates a greater degree of cultural awareness within that domain. The reliability was 0.91 for students and 0.82 for faculty. The subscales' reliability values ranged from 0.66 to 0.99 for students and from 0.56 to 0.87 for faculty [22].

MAIRS-MS

The MAIRS-MS instrument, initially created by Karaca, has demonstrated its validity and applicability in broader healthcare education settings, including nursing [23, 24]. It requires participants to rate their self-assessment on a 5-point Likert scale (ranging from 1 = 'strongly disagree' to 5 = 'strongly agree') on 22 statements regarding their AI readiness [25]. The MAIRS-MS is subdivided into four factors: cognition (8 items, score range: 8–40), ability (8 items, score range: 8–40), vision (3 items, score range: 3–15), and ethics (3 items, score range: 3–15). Total score ranges from 22 to 110. Higher scores on the total scale or any subscale reflect greater readiness for the integration of artificial intelligence in medical education and practice. Two exemplary items are 'I can explain how AI systems are trained' (cognition) and 'I can explain the limitations of AI technology' (vision) [25]. The internal consistency of the overall scale was acceptable (Cronbach's alpha = 0.88), and the Cronbach's alphas for the individual factors were 0.83 (cognition), 0.77 (ability), 0.72 (vision), and 0.63 (ethics) [26].

SET-M

The participants' confidence was measured using the SET-M at T1 only. The modified version of this instrument was published in 2015, with a total of 19 items and a 3-point response scale (ranging from 1 = 'do not agree' to 3 = 'strongly agree') [27]. The items are divided into four domains: pre-briefing (2 items, score range: 2–6), learning (6 items, score range: 6–18), confidence (6 items, score range: 6–18), and debriefing (5 items, score range: 5–15). Total score ranges from 19 to 57. A higher score on the total scale or any subscale indicates a more positive perception of the effectiveness of simulation-based learning. Cronbach's alpha for the overall SET-M was 0.94, with the reliability of each domain ranging from 0.83 to 0.91 [27].

Demographic questionnaire

A questionnaire was used to collect the participants' demographic data, including gender, year of study, and previous clinical experience, such as Temporary Undergraduate Nursing Students (TUNS). TUNS are nursing students who work part-time in healthcare settings while completing their nursing education. Specifically, they are typically students in their 4th or 5th year of a nursing degree program.

Both A and B completed a three-day simulation intervention, with a one-week washout interval implemented between intervention phases. This washout period—during which no study interventions were administered—was designed to minimize carryover effects and allow any residual learning effects from the first intervention to dissipate before the second intervention began, consistent with prior educational research [28]. To further reduce carryover effects, different clinical scenarios were used at each time point: participants experienced the acute pneumonia case during the first intervention phase (T1) and the acute appendicitis case during the second intervention phase (T2), regardless of which simulation modality they received. Five questionnaire tools were used to collect data: the Clinical Competence Questionnaire (CCQ), the Cultural Awareness Scale (CAS), the Medical Artificial Intelligence Readiness Scale for Medical Students (MAIRS-MS), the Simulation Effectiveness Tool – Modified (SET-M), and a demographic questionnaire. The questionnaires were administered to the participants at three time points – pre-intervention (baseline [T0]), time 1 (T1), and time 2 (cross-over session) (T2).

Intervention details

The participants engaged in two distinct clinical scenarios in person: an acute pneumonia case at T1 and an acute appendicitis case at T2. At both T1 and T2, 35They participated in the same structured debrThey participated in the same structured debrThey participated in the same structured debrThey participated in the same structured debrThey participated in the same structured debrThey participated in the same structured debrThey participated in the same structured debrThey participated in the same structured debrThey participated in the same structured debrThey participated in the same structured debrThey participated in the same structured debrThey participated in the same structured debriefing protocol utilizing the 3D debriefing model [29], using the same set of debrief questions and facilitated by trained staff to ensure consistency across all sites. However, the type of simulation differed between the groups: one group experienced the GenAI patient simulation while the other group participated in the 360 VR simulation. Each intervention lasted 60 min, followed by a 60-minute, face-to-face debriefing session. After the first round, the groups crossed over so that all students experienced both types of simulation. This design enabled a fair comparison of educational outcomes while ensuring that all students received equal exposure to both simulation modalities and clinical scenarios.

Description of the two clinical case interventions

360° VR simulation - case 1: acute pneumonia

The intervention for Case 1 with 360° VR simulation, focusing on a 25-year-old patient named Peter, who had been diagnosed with acute pneumonia. The session was structured as follows:

Pre-briefing: Introduction to learning objectives and orientation to the LMS platform for interactive video engagement.

Case Background: Overview of Peter's clinical presentation and relevant history.

Activity 1: Students watched the initial video segment, identified and timestamped health problems, and engaged in cultural discussions about stress coping and nursing practices.

Activity 2: The following video segment focused on physical assessment, with students identifying and timestamping assessment findings, discussing missed assessments, and sharing cultural perspectives on care.

Activity 3: Group discussion on prioritizing three nursing diagnoses, referencing current feedback and North American Nursing Diagnosis Association International (NANDA-I) [30].

Activity 4: The final video segment involved group discussion on nursing interventions and patient education, focusing on evidence-based practices for airway clearance, gas exchange, hyperthermia, nutrition, and fluid balance. Written summaries and rationales were provided for each nursing diagnosis.

Generative AI patient Simulation- case: acute pneumonia

The content and case scenario were identical to Group A, but students interacted with an AI-simulated patient. They asked questions related to history-taking and physical assessment, then formulated nursing diagnoses based on the AI's responses. Clinical reasoning was required to determine appropriate nursing diagnoses and interventions, with all activities mirroring the VR simulation structure (pre-briefing, case background, assessment, diagnosis, and intervention planning). The interaction was dynamic and student-led, guided by their inquiries and the AI's generated responses.

360° VR simulation - case 2: acute appendicitis

The intervention for Case 2 with 360° VR simulation centered on a 38-year-old patient, Ken Chan, admitted for acute appendicitis and post-emergency appendectomy. The session was structured as follows:

Pre-briefing: Explanation of learning objectives (nursing assessment, prioritization of nursing diagnoses, cultural differences in care) and orientation to the LMS platform.

Case Background: Introduction to Ken's clinical scenario, including his admission for acute lower abdominal pain, vomiting, and subsequent emergency surgery. Relevant history included high job stress, smoking, alcohol use, and initial signs of liver cirrhosis.

Activity 1: Students watched the initial video segment, identified and timestamped health problems, and participated in cultural discussions about health-seeking behaviors, healthcare systems, and pain management.

Activity 2: The following video segment focused on physical assessment, with students identifying and timestamping assessment findings (e.g., wound assessment, pain level, vital signs, substance use), discussing missed assessments, and sharing cultural perspectives on pain assessment and wound healing.

Activity 3: Group discussion to identify and prioritize three nursing diagnoses (e.g., acute pain, risk for deficient fluid volume, risk for ineffective breathing patterns), referencing expert feedback and NANDA-I guidelines.

Activity 4: The final video segment involved group discussion on nursing interventions and patient education, emphasizing evidence-based care for pain, fluid balance, infection prevention, nutrition, and liver function. Cultural sharing included discussion of clinical practices, anti-smoking and alcohol cessation policies, and relevant health apps.

Generative AI patient simulation- case 2: acute appendicitis

The same content for the acute appendicitis case was used, but students interacted with an AI patient. They conducted history-taking and physical assessments by asking the AI questions, then formulated nursing diagnoses and planned interventions based on the AI's responses. The process followed the same structure as the VR simulation (pre-briefing, case background, assessment, diagnosis, and intervention planning), but the interaction was dynamic and student-driven, with clinical reasoning applied to each AI-generated response. No results or outcomes are reported, as per instructions.

Note: In both cases, the only difference between Group A and Group B is the mode of simulation (360° VR video vs. Generative AI patient interaction); the clinical content, learning objectives, and session structure remain consistent across groups.

Components of GenAI patient simulation

The proposed GenAI Patient Simulation system for nursing education integrates several key components to enhance the learning experience. The core educational innovation lies in our pedagogical framework rather than the specific AI technology. The GenAI system utilizes a specialized Large Language Model (LLM) developed through collaboration with 0xmd Inc. (Hong Kong). This LLM was purpose-built through proprietary development and trained on a comprehensive corpus comprising medical datasets, real-world clinical data, academic research papers, and peer-reviewed literature to ensure high domain-specific accuracy and robust clinical reasoning capabilities. While the specific model is proprietary, our complete educational framework is reproducible through: (1) detailed clinical case protocols with specific patient presentations and assessment criteria, (2) structured NANDA-I-based evaluation rubrics, (3) comprehensive debriefing protocols using the 3D model, and (4) scoring algorithms. Researchers can implement this pedagogical approach using any conversational AI platform by following our documented framework, ensuring educational reproducibility regardless of the underlying AI technology.

Unlike generic LLMs, it is optimized explicitly for medical and healthcare applications. The School of Nursing at the University of Hong Kong collaborated with the company to adapt and fine-tune the model for nursing education, incorporating clinical simulation cases, structured diagnostic frameworks, and pedagogical methodologies. This co-development enables realistic, culturally relevant, and educationally aligned patient interactions within the simulation system, supporting its use as an innovative tool for enhancing clinical reasoning and decision-making skills in nursing education.

The system facilitates structured learning through an interactive dialogue interface where students systematically engage with AI patients. Students are guided to conduct comprehensive history-taking using open-ended questions, allowing them to gather essential patient information while developing their communication skills.

The interface prompts students to perform virtual physical assessments by requesting specific examinations, enabling them to practice clinical assessment techniques in a risk-free environment. Based on the information gathered, students are challenged to formulate differential diagnoses, encouraging critical thinking and analytical reasoning skills essential for nursing practice. Furthermore, the system requires students to propose evidence-based interventions with clear rationales, fostering the development of clinical decision-making abilities grounded in current best practices.

The AI patient component responds consistently with programmed clinical presentations, providing realistic answers that accurately reflect the patient's cultural background, symptom severity, and emotional state, all of which are appropriate to the specific medical condition. This consistency ensures that all students encounter standardized scenarios while maintaining the authenticity of patient interactions. The system categorizes nursing diagnoses into distinct domains such as physiological, psychosocial, and cultural, further classifying them into actual, potential, and wellness diagnoses to provide a structured approach to patient care assessment. Maslow's hierarchy of needs was explicitly used to train the AI system to recognize and prioritize patient care objectives, enabling the AI to evaluate student responses, generate contextually relevant debriefing feedback, and assign scores based on how well students address care priorities, such as managing physiological needs before addressing psychosocial concerns. Additionally, the system employs a standardized scale to assess the severity and complexity of disease conditions, considering factors like symptom acuity, potential for rapid deterioration, and intervention urgency, which aids students in identifying critical nursing diagnoses. Finally, the system establishes care priorities through a scoring mechanism that evaluates each diagnosis based on its impact on patient safety, potential complications, and overall outcomes, ensuring informed clinical decision-making.

Implementation of constructive feedback through GenAI debriefing

The implementation of constructive feedback through GenAI debriefing in the proposed system is designed to facilitate asynchronous learning by providing students with personalized, post simulation expert feedback. This process involves reflective practice where students analyze their responses, actions, and the outcomes of their clinical decisions. GenAI evaluates the rationale behind each nursing diagnosis, assessing the alignment with established grading frameworks to highlight areas of strength and improvement, thereby deepening students' understanding of prioritization. Additionally, it scrutinizes the actions taken during simulations, offering feedback on the appropriateness of interventions to enhance clinical judgment and critical thinking. Finally, GenAI assesses the results of these actions, providing insights into the effectiveness of interventions based on chosen diagnoses, reinforcing evidence-based practice. This debriefing not only offers expert feedback on the correct prioritization of nursing diagnoses but also generates a score based on the accuracy of assessments and choices, ensuring a comprehensive learning experience.

Implementation of 360° VR simulation

The 360° VR content was captured using the Insta360 Pro 2 camera system, which provides high-resolution 360-degree video recording capabilities suitable for creating immersive educational content. The raw 360° footage was processed and stitched using the manufacturer's software to create seamless panoramic videos. These processed videos were then uploaded and hosted on the institution's LMS, ensuring consistent delivery across all participating sites. The use of standardized hardware and processing workflows facilitated replication across the six international locations while maintaining consistency in video quality and immersive experience. The 360° VR simulation employed in this study was developed in-house by the research team at the University of Hong Kong. This system is designed for ease of replication in diverse educational settings, requiring only standard 360-degree cameras for content creation and compatible playback devices (e.g., computers, tablets, or VR headsets) for delivery. Educators can readily adapt this approach by producing customized 360° video scenarios that align with specific learning objectives relevant to their curriculum.

For this study, students assigned to the 360° VR simulation group viewed immersive video scenarios via a screen interface. The video content was identical to that used in the GenAI group, ensuring consistency across experimental conditions. The immersive nature of the 360° VR experience facilitated active decision-making and interactive engagement, as students could freely navigate the environment by scrolling within the video using a mouse to observe the complete 360-degree perspective.

The 360° VR videos were hosted on the institution's LMS [31], which supported interactive features, including timestamped commenting. Students were provided with a structured question guide to direct their observations and reflections as they engaged with the scenarios. This allowed students to annotate specific moments within the video by placing timestamped comments, thereby capturing their insights and questions in real-time.

Following the simulation, a faculty-led debriefing session was conducted. During this session, students reviewed and discussed each other's timestamped comments, fostering vicarious peer learning and reflective

Table 1 Baseline characteristic of the participants ($n=44$)

	Participants, No (%)			
Characteristic	**All (*N*=44)**	**A group (*N*=22)**	**B group (*N*=22)**	***P* value**
Age, mean (SD)	21.0 (1.24)	21.1 (1.27)	20.9 (1.23)	0.632
Sex				0.216
Male	7 (38.3)	5 (22.7)	2 (9.1)	
Female	37 (61.7)	17 (77.3)	20 (90.9)	
Year of study				0.627
Year 1	2 (4.5)	1 (4.5)	1 (4.5)	
Year 2	15 (34.1)	6 (27.3)	9 (40.9)	
Year 3	27 (61.4)	15 (68.2)	12 (54.5)	
Do you have any TUNS experience?				0.262
Yes	9 (20.5)	3 (13.6)	6 (27.3)	
No	35 (79.5)	19 (86.4)	16 (72.7)	
Have you attended any clinical practicum?				1.000
Yes	28 (73.6)	14 (73.6)	14 (73.6)	
No	8 (36.4)	8 (36.4)	8 (36.4)	
Do you have any GenAI training before?				0.240
Yes	8 (18.2)	6 (27.3)	2 (9.1)	
No	36 (81.8)	16 (72.7)	20 (90.9)	
CCQ Total Score, mean (SD)	154.16 (32.82)	157.23 (31.38)	151.09 (34.66)	0.542
Nursing Professional Behaviors, mean (SD)	52.25 (13.99)	52.82 (13.97)	51.68 (14.33)	0.791
General Performance, mean (SD)	46.50 (6.40)	46.68 (5.97)	46.32 (6.94)	0.853
Core Nursing Skills, mean (SD)	37.91 (11.20)	39.18 (11.23)	36.64 (11.28)	0.457
Advanced Nursing Skills, mean (SD)	17.50 (5.50)	18.55 (4.90)	16.45 (5.97)	0.211
CAS Total Score, mean (SD)	189.43 (40.15)	190.91 (42.38)	187.95 (38.74)	0.810
General Educational Experience	74.75 (16.69)	74.00 (17.46)	75.50 (16.27)	0.770
Cognitive Awareness	38.86 (8.42)	38.05 (8.38)	39.68 (8.58)	0.526
Research Issue	20.77 (5.17)	20.41 (5.32)	21.14 (5.10)	0.646
Behaviors or Comfort with Interactions	25.73 (8.71)	27.36 (9.08)	24.09 (8.19)	0.216
Patient Care or Clinical Issues	29.32 (7.69)	31.09 (7.76)	27.55 (7.38)	0.128
MAISMS Total Score, mean (SD)	60.27 (14.45)	61.00 (17.41)	59.55 (11.10)	0.743
Cognition	22.30 (6.11)	22.77 (7.08)	21.82 (5.09)	0.610
Ability	21.77 (5.85)	21.86 (6.65)	21.68 (5.07)	0.919
Vision	8.05 (2.40)	8.27 (2.68)	7.82 (2.13)	0.536
Ethics	8.16 (2.78)	8.09 (2.78)	8.23 (2.84)	0.873

practice. The LMS platform's structured debriefing feature provided educators with a systematic framework to guide the discussion, ensuring that feedback was timely, relevant, and closely aligned with the learning objectives of the simulation.

The high degree of realism afforded by the 360° VR scenarios enhanced the authenticity of the learning experience, supporting the educational aims of the intervention. The described implementation is readily transferable to other research and educational contexts, provided that the necessary technical infrastructure is available. This framework enables other researchers and educators to replicate the intervention, customize content, and systematically evaluate student engagement and learning outcomes using immersive 360° VR technology.

Implementation of debriefing in both groups

Both the 360 VR and GenAI patient simulation groups participated in structured debriefing sessions following their simulation experiences. For both groups, accredited simulation educators and six core faculty members experienced in simulation-based education conducted debriefings using the 3D Model, which encompasses three phases: Defusing (emotional release), Discovering (performance analysis), and Deepening (application of learning). All facilitators underwent comprehensive training in the 3D debriefing framework to ensure consistency and quality across sessions.

In the GenAI patient simulation group, participants received additional, individualized constructive feedback generated through GenAI-driven debriefing. This AI-powered feedback provided tailored insights and recommendations based on each learner's performance, further enriching the reflective learning process.

By combining facilitator-led debriefing with GenAI-generated feedback (for the GenAI group), both groups benefited from a robust, evidence-based approach that supports reflective learning, clinical decision-making, and emotional engagement, fully aligned with best practices in simulation-based education.

Data analysis

Table 1 shows the descriptive statistics, including means and standard deviations, were used to summarise the participants' demographics and scores on the CCQ, CAS, MAIRS-MS, and SET-M. The baseline characteristics of the intervention and control groups were compared using chi-square tests or Fisher's exact tests for categorical data and independent t-tests for continuous data.

A linear mixed-model analysis was conducted to analyse the CCQ, CAS, and MAIRS-MS scores. This model treated the participants as a random effect and included time, group, randomisation sequence order, and the time–group interaction as fixed effects [32]. Linear contrasts were used to assess both between-group differences and within-group changes over time. All of the statistical analyses were performed using IBM SPSS version 29.0 (IBM Corp., Armonk, NY, USA). Two-sided tests were used throughout, with a significance level of $p < 0.05$.

Table 2 Simulation effectiveness Tool-Modified (SET-M) responses

Simulation Effectiveness Tool-Modified (SET- M) responses All respondents (*N* = 44)			
Survey questions:	**Strongly Agree**	**Somewhat Agree**	**Do Not Agree**
Prebriefing subscale			
Prebriefing increased my confidence.	52.3% (23)	47.7% (21)	0% (0)
Prebriefing was beneficial to my learning.	65.9% (29)	34.1% (15)	0% (0)
Learning subscale			
I am better prepared to respond to changes in my patient's condition.	70.5% (31)	29.5% (13)	0% (0)
I developed a better understanding of the pathophysiology.	68.2% (30)	31.8% (14)	0% (0)
I am more confident of my nursing assessment skills.	68.2% (30)	31.8% (14)	0% (0)
I felt empowered to make clinical decisions.	77.3% (34)	22.7% (10)	0% (0)
I developed a better understanding of medications.	65.9% (29)	31.8% (14)	2.3% (1)
I had the opportunity to practice my clinical decision making skills.	70.5% (31)	29.5% (13)	0% (0)
Confidence subscale			
I am more confident in my ability to prioritize care and interventions	68.2% (30)	31.8% (14)	0% (0)
I am more confident in communicating with my patient.	72.7% (32)	27.3% (12)	0% (0)
I am more confident in my ability to teach patients about their illness and interventions.	70.5% (31)	29.5% (13)	0% (0)
I am more confident in my ability to report information to health care team.	68.2% (30)	31.8% (14)	0% (0)
I am more confident in providing interventions that foster patient safety.	75.0% (33)	25.0% (11)	0% (0)
I am more confident in using evidence-based practice to provide nursing care.	72.7% (32)	27.3% (12)	0% (0)
Debriefing subscale			
Debriefing contributed to my learning.	75.0% (33)	25.0% (11)	0% (0)
Debriefing allowed me to verbalize my feelings before focusing on the scenario.	68.2% (30)	31.8% (14)	0% (0)
Debriefing was valuable in helping me improve my clinical judgment.	68.2% (30)	31.8% (14)	0% (0)
Debriefing provided opportunities to self-reflect on my performance during simulation.	68.2% (30)	31.8% (14)	0% (0)
Debriefing was a constructive evaluation of the simulation.	75.0% (33)	25.0% (11)	0% (0)

Results

Descriptive results

In total, 44 participants were included in the study, divided equally into two groups: Group A ($n = 22$) and Group B ($n = 22$). The mean age of participants was 21.0 ± 1.24 years, with no significant difference between groups ($P = 0.632$). Most participants were female (84.1%), and the distribution of sex did not differ significantly between groups ($P = 0.216$). Regarding the year of study, most participants were in Year 2 (34.1%), with no significant difference between groups ($P = 0.627$).

Participants with prior TUNS experience comprised 20.5%, while those without accounted for 79.5%, with no significant difference between groups ($P = 0.262$). Similarly, the majority (73.6%) had attended clinical practicum, and this characteristic was evenly distributed across groups ($P = 1.000$). Only a minority (18.2%) had experience with GenAI training, with no significant difference between groups ($P = 0.240$).

Baseline scores for CCQ Total Score, Nursing Professional Behaviors, General Performance, Core Nursing Skills, and Advanced Nursing Skills were comparable between groups, with P values ranging from 0.211 to 0.853, indicating no significant differences. The CAS Total Score also showed no significant difference between groups ($P = 0.810$). These findings suggest that demographic and baseline clinical characteristics were well-balanced across the two groups ($P > 0.05$).

Effect of the interventions on SET-M

Table 2 shows the responses of the participants to the SET-M. Most of the participants (75%, $n = 33$) strongly agreed that debriefing contributed to their learning. Regarding confidence subscale, over half proportion of participants (68.2%, $n = 30$) strongly agreed that they are more confident of their nursing assessment skills. Additionally, most participants (72.7%, $n = 32$) strongly agreed that they were confident in using evidence-based practice to provide nursing care.

Effects of the interventions on CCQ

Both groups demonstrated significant within-group improvements in perceived clinical competence at T1 and T2. Group B (GenAI first) showed significantly greater improvement in CCQ total score at T1 compared to Group A, with a between-group difference of 16.59 points (95% CI: 2.77, 30.41, $p = 0.020$). By T2, both groups maintained significant improvements from baseline, suggesting sustained educational benefits across the crossover design.

Effects of the interventions on CAS

Both groups demonstrated improvements in cultural awareness. Group B showed consistent significant gains at both time points [T1: 27.59 (95% CI: 9.54, 45.64), $p = 0.003$; T2: 20.91 (95% CI: 2.86, 38.96), $p = 0.024$], while Group A showed significant improvement at T1 [18.27 (95% CI: 0.22, 36.32), $p = 0.047$] with marginal significance at T2. No significant between-group differences were detected for total scores or individual subscales, indicating both interventions effectively enhanced culturalawareness.

Effects of the interventions on MAIRS-MS

The MAIRS-MS was administered at baseline (T0) and following each group's GenAI simulation exposure, as this instrument specifically measures nursing students' AI readiness after exposure to AI technology. Given the cross-over design, Group A received MAIRS-MS at T0 and T2 (following their GenAI exposure at T2), while Group B received MAIRS-MS at T0 and T1 (following their GenAI exposure at T1). No MAIRS-MS measurement was conducted following the 360° VR simulation, as this tool is designed to assess AI readiness specifically after exposure to AI intervention.

Table 3 presents the results of mixed-effects tests on the interventions' effects on the MAIRS-MS. For the MAIRS-MS Total Score, Group A exhibited significant changes from baseline (T0) to post-GenAI exposure (T2) with an increase of 16.64 (95% CI [9.80, 23.47], $p < 0.001$). In comparison, Group B showed significant changes from baseline (T0) to post-GenAI exposure (T1) with an increase of 30.18 (95% CI [23.35, 37.01], $p < 0.001$). A significant between-group difference was noted (12.09, 95% CI [4.43, 19.75], $p = 0.003$), with Group B demonstrating greater improvement in AI readiness. Significant between-group differences were also observed in the Cognition and Ability subscales. Overall, both groups showed significant improvements in AI readiness following their respective GenAI exposures.

Discussion

This study examined the impact of two innovative interventions – GenAI and 360° VR simulation – on the perceived clinical competence, cultural awareness, and AI readiness of nursing students in two distinct groups. Group A received the 360° VR Simulation first, followed by GenAI, while Group B experienced these interventions in the reverse order. Both groups demonstrated significant improvements in perceived clinical competence, with Group B initially showing advantages in general performance and core nursing skills, possibly due to the adaptive and interactive nature of GenAI simulation. Regarding cultural awareness, Group B made earlier gains in educational experience and cognitive awareness, although Group A eventually matched these improvements. Notably, both groups exhibited substantial progress in AI readiness across all subscales.

Additionally, a majority of students reported positive experiences with debriefing, contributing to increased confidence in their nursing skills. Both 360° VR simulation and GenAI patient simulation appear to be promising pedagogical strategies for enhancing certain clinical outcomes among the nursing students in this study. However, GenAI showed a trend toward a more pronounced impact in this cohort, possibly due to higher initial engagement, although further research is needed to confirm these findings. Specifically, when GenAI patient simulation is provided first, it yields greater outcomes; subsequently, utilising 360° VR can enhance these results.

Clinical competence

The study's results indicate that the intervention enhanced perceived clinical competence in both groups, with Group B demonstrating a more pronounced initial improvement at T1, particularly in the General Practice (GP) and Clinical Nursing Skills (CNS) components of the Clinical Competence Questionnaire (CCQ). However, by T2, Group A had begun to catch up, suggesting that while their progress was initially slower, they eventually reached comparable levels of competence, which may indicate that the intervention had a more gradual impact on some participants but was ultimately effective for both groups over time. The customisation enabled by GenAI interventions, which catered to individual learners' needs, preferences, and learning styles, likely facilitated targeted feedback, adaptive challenges, and content alignment with specific clinical roles and responsibilities. Our findings demonstrated that a tailored, personalised learning approach—developed and implemented in our study—effectively fosters essential clinical competencies among nursing students. This aligns with existing literature, which also supports the effectiveness of personalised learning in enhancing clinical competence across both generalist and specialist nursing practice [33]. Moreover, GenAI interventions specifically targeted cognitive skills such as problem-solving, critical thinking, decision-making, and clinical reasoning, which are crucial for effective practice in healthcare settings. Engaging participants in higher-order thinking tasks and complex scenarios relevant to their practice areas potentially enhanced the development of these cognitive skills more effectively than traditional interventions.

A notable feature of GenAI is its ability to deliver real-time feedback, debriefing, performance monitoring, and adaptive learning pathways tailored to learners' responses and progress. This immediate feedback loop enables participants to pinpoint areas for improvement, adjust their strategies, and track their progress over time, thereby fostering an environment conducive to continuous learning and skill refinement. The personalised feedback and adaptive nature of GenAI likely contributed to

Table 3 Mixed effects analysis for the interventional effects ($n = 44$)

Measure	Group A($n = 22$) Within-Group Change (95% CI)	*P* Value	Group B($n = 22$) Within-Group Change (95% CI)	*P* Value	Between-Group Difference Mean (95% CI)	*P* Value
CCQ-Total						
T1	24.95 (13.96, 35.95)	< 0.001*	47.68 (36.68, 58.68)	< 0.001*	16.59 (2.77, 30.41)	0.020*
T2	31.09 (20.09, 42.09)	< 0.001*	39.64 (28.64, 50.64)	< 0.001*	2.41 (-11.41, 16.23)	0.727
CCQ-NP						
T1	8.82 (3.70, 13.94)	< 0.001*	11.86 (6.74, 16.98)	< 0.001*	1.91 (-5.64, 9.45)	0.612
T2	12.50 (7.38, 17.62)	< 0.001*	13.59 (8.47, 18.71)	< 0.001*	-0.05 (-7.59, 7.50)	0.990
CCQ-GP						
T1	3.36 (0.45, 6.28)	0.024*	9.41 (6.49, 12.33)	< 0.001*	5.68 (2.72, 8.64)	< 0.001*
T2	3.73 (0.81, 6.65)	0.013*	6.77 (3.85, 9.69)	< 0.001*	2.68 (-0.28, 5.64)	0.075
CCQ-CNS						
T1	9.23 (5.06, 13.40)	< 0.001*	18.14 (13.96, 22.31)	< 0.001*	6.36 (2.08, 10.65)	0.005*
T2	9.36 (5.19, 13.54)	< 0.001*	13.82 (9.65, 17.99)	< 0.001*	1.91 (-2.38, 6.20)	0.374
CCQ-ANS						
T1	3.55 (1.31, 5.78)	0.002*	8.27 (6.04, 10.51)	< 0.001*	2.64 (-0.39, 5.66)	0.086
T2	5.50 (3.27, 7.73)	< 0.001*	5.45 (3.22, 7.69)	< 0.001*	-2.14 (-5.16, 0.89)	0.162
CAS-Total						
T1	18.27 (0.22, 36.32)	0.047*	27.59 (9.54, 45.64)	0.003*	6.36 (-15.08, 27.80)	0.552
T2	15.32 (-2.73, 33.37)	0.095	20.91 (2.86, 38.96)	0.024*	2.64 (-18.80, 24.08)	0.805
CAS-GEE						
T1	3.95 (-3.92, 11.83)	0.321	5.41 (-2.46, 13.28)	0.175	2.95 (-6.70, 12.61)	0.540
T2	3.32 (-4.55, 11.19)	0.404	2.50 (-5.37, 10.37)	0.529	0.68 (-8.97, 10.34)	0.887
CAS-CA						
T1	4.77 (1.02, 8.53)	0.013*	6.09 (2.33, 9.85)	0.002*	2.95 (-1.01, 6.92)	0.140
T2	4.23 (0.47, 7.98)	0.028*	6.05 (2.29, 9.80)	0.002*	3.45 (-0.51, 7.42)	0.086
CAS-RI						
T1	2.68 (-0.08, 5.44)	0.057	2.05 (-0.72, 4.81)	0.145	0.09 (-3.17, 3.36)	0.955
T2	0.64 (-2.13, 3.40)	0.648	1.50 (-1.26, 4.26)	0.283	1.59 (-1.67, 4.86)	0.331
CAS-BOCWI						
T1	2.86 (-1.16, 6.89)	0.161	5.59 (1.57, 9.62)	0.007*	-0.55 (-5.43, 4.34)	0.823
T2	2.68 (-1.34, 6.71)	0.189	4.00 (-0.03, 8.03)	0.051	-1.95 (-6.84, 2.93)	0.424
CAS-PCOCI						
T1	4.00 (0.55, 7.45)	0.023*	8.45 (5.01, 11.90)	< 0.001*	0.91 (-3.00, 4.82)	0.641
T2	4.45 (1.01, 7.90)	0.012*	6.86 (3.42, 10.31)	< 0.001*	-1.14 (-5.04, 2.77)	0.560
MAIRS-MS-Total						
(T2 for Group A, T1 for group B)	16.64 (9.80, 23.47)	< 0.001*	30.18 (23.35, 37.01)	< 0.001*	12.09 (4.43, 19.75)	0.003*
MAIRS-MS-Cognition						
(T2 for Group A, T1 for group B)	4.50 (1.46, 7.54)	0.005*	11.59 (8.55, 14.63)	< 0.001*	6.14 (2.89, 9.38)	< 0.001*
MAIRS-MS-Ability						
(T2 for Group A, T1 for group B)	6.00 (2.25, 9.75)	0.002*	10.36 (6.62, 14.11)	< 0.001*	4.18 (0.13, 8.23)	0.043*
MAIRS-MS-Vision						
(T2 for Group A, T1 for group B)	3.23 (1.79, 4.66)	< 0.001*	4.55 (3.11, 5.98)	< 0.001*	0.86 (-0.60, 2.33)	0.241

Table 3 (continued)

Measure	Group A($n=22$) Within-Group Change (95% CI)	*P* Value	Group B($n=22$) Within-Group Change (95% CI)	*P* Value	Between-Group Difference Mean (95% CI)	*P* Value
MAIRS-MS-Ethics						
(T2 for Group A, T1 for group B)	2.91 (1.63, 4.19)	< 0.001*	3.68 (2.40, 4.96)	< 0.001*	0.91 (-0.53, 2.35)	0.209

Note:

• TUNS: Temporary Undergraduate Nursing Student

• CCQ: Clinical Competence Questionnaire

• NP: Nursing professional behaviors

• GP: General Performance

• CNS: Core Nursing Skills

• ANS: Advanced Nursing Skills

• CAS: Cultural Awareness Scale

• GEE: General Educational Experience

• CA: Cognitive Awareness

• RI: Research Issues

• BOCWI: Behavior or Comfort with Interactions

• PCOCI: Patient Care or Clinical Issues

*$p < 0.05$ indicates statistical significance

the improved performance outcomes observed in the GP and CNS assessments of the CCQ. By leveraging advanced AI technologies, including deep learning and natural language processing, GenAI creates dynamic, responsive, and intelligent learning environments that optimize educational experiences, customize content delivery, and support individualized learning pathways focused on the specific competencies required for diverse nursing specialties.

Cultural awareness

The interventions aimed at enhancing cultural awareness, as measured by the Cultural Awareness Scale (CAS), demonstrated significant improvements in both groups, particularly in the dimensions of Cognitive Awareness and Patient Care or Clinical Issues. Group B exhibited more substantial enhancements at both time points compared to Group A, indicating a potentially more effective intervention for this group. Nonetheless, both groups experienced significant gains in several core areas of cultural awareness.

The notable advancements in Cognitive Awareness and patient care, as well as clinical issues, underscore the efficacy of these interventions in improving participants' understanding of cultural issues and their practical application in clinical settings. This is further substantiated by evidence in the literature that cultural awareness is increasingly recognized as vital for delivering high-quality and patient-centered care [34, 35]. The balanced representation of students from various countries in both groups facilitated an increase in cultural awareness through both patient simulations. These simulations were designed to interact according to distinct cultural beliefs and express cultural requirements, allowing students to engage in and gain insights from culturally nuanced exchanges, thereby enhancing their comprehension of cultural diversity within healthcare environments. The improvements in these areas suggest that the intervention successfully addressed key components of cultural awareness, potentially leading to better patient outcomes and more culturally sensitive care practices.

However, the absence of significant changes in the areas of Behavior or Comfort with Interactions and Research Issues across both groups highlights a critical gap in the interventions. These components of cultural awareness, which involve interpersonal dynamics and behavior in diverse settings, may necessitate more teaching sessions, such as role-playing, simulations, or real-world applications, to cultivate practical skills in navigating cultural interactions. The absence of significant changes in the Behavior or Comfort with Interactions and Research Issues subscales may reflect several methodological considerations. First, our interventions were primarily designed to enhance clinical reasoning and decision-making rather than interpersonal cultural interaction skills. Second, the intervention duration (60 min per simulation) and short study timeframe may have been insufficient to impact these behavioral dimensions of cultural awareness. Additionally, these subscales may require measurement approaches specifically designed to detect changes following educational interventions. Future research should investigate whether interventions specifically targeting interpersonal cultural skills, with appropriate measurement tools and extended follow-up periods, might be more effective for addressing these domains of cultural awareness.

The findings indicate that interventions focusing on cultural awareness may enhance healthcare professionals' cognitive understanding and clinical application of cultural awareness. Such improvements are essential for fostering culturally sensitive healthcare environments, particularly in increasingly diverse patient populations. However, to achieve broader cultural awareness, future training programs should incorporate more specific elements related to research issues and intercultural interactions, areas that were not significantly addressed in this study.

AI readiness

The findings indicate that GenAI patient simulation was effective in enhancing AI readiness for both groups, with both demonstrating significant improvements from their respective baselines after exposure to GenAI. Group B showed a larger improvement [30.18 (95% CI: 23.35, 37.01), $p < 0.001$] compared to Group A [16.64 (95% CI: 9.80, 23.47), $p < 0.001$], with a significant between-group difference [12.09 (95% CI: 4.43, 19.75), $p = 0.003$]. Notably, Group B was assessed at T1 immediately following their first GenAI exposure, while Group A was assessed at T2 following their GenAI exposure after experiencing a 360° VR simulation first.

The significant improvements observed across all subscales of the MAIRS-MS suggest that GenAI patient simulation was highly effective in increasing nursing students' readiness to engage with and utilize AI in their practice. The larger improvement in Group B may reflect the immediate impact of GenAI simulation when experienced as the first intervention, before any potential carry-over effects from prior simulation exposure. This sequence effect suggests that initial exposure to GenAI simulation may optimize learning gains in AI readiness, particularly when students encounter AI technology without prior simulation experience in the study context.

The results indicate that GenAI patient simulation successfully addressed key aspects of AI readiness, including cognitive understanding of AI concepts, the ability to interact with AI systems, and vision for AI integration in healthcare. While our study measured ethical readiness as part of the MAIRS-MS instrument, further analysis of specific ethical considerations would strengthen future research in this domain. The substantial increases in scores, particularly in the Cognition and Ability subscales, highlight that participants gained both theoretical knowledge about AI and practical skills in working with AI technologies through direct interaction with the GenAI patient simulation system. This comprehensive improvement across all dimensions of AI readiness is crucial for preparing nursing students to implement AI in healthcare settings successfully.

Perceived simulation effectiveness (SET-M)

The analysis of the SET-M responses indicates that a majority of the participants perceived debriefing as highly beneficial to their learning, with 75% expressing strong agreement. Furthermore, a large majority of the participants reported increased confidence in their nursing assessment skills (68.2%). These findings underscore the importance of perceived competence and confidence in clinical training, suggesting that structured peer interactions may enhance learning experiences and foster a supportive educational environment.

Implications of nursing educational practice

The integration of GenAI patient simulation and immersive 360° VR may offer innovative approaches to developing clinical competence, cultural awareness, and AI readiness among undergraduate nursing students. Both modalities provide realistic and interactive environments that enhance experiential learning and clinical decision-making, thereby addressing the limitations of traditional didactic methods. GenAI enables adaptive, personalized feedback and real-time debriefing, supporting individualized learning trajectories and fostering critical thinking. These innovations may help equip future nurses with relevant skills and confidence; further research is needed to determine their impact on patient care and safety.

Implications of the findings

The findings suggest that both GenAI patient simulation and 360° VR simulation may improve participants' perceived clinical competence, cultural awareness, and AI readiness, with GenAI showing a greater effect in some measured domains. The positive student feedback on debriefing and increased confidence in assessment skills highlight the pedagogical value of simulation-based education. These results provide preliminary support for incorporating advanced simulation technologies into nursing curricula to enhance readiness for AI-integrated healthcare; however, further large-scale studies are warranted to confirm this finding.

Future directions

Future research should investigate longitudinal follow-up at 3, 6, and 12 months after GenAI and 360° VR simulation to assess the impact on clinical competence. Qualitative investigations into student experiences can deepen understanding of the mechanisms driving observed improvements. Expanding outcome measures to include patient care outcomes and the transferability of skills to real-world settings will further inform the development of the curriculum. Objective assessments and culturally responsive models should be incorporated to address the limitations of self-report tools and better capture diversity in learning needs. Additionally, future studies should consider larger sample sizes ($n \geq 100$) per group, more diverse samples to enable subgroup analyses by country and context, enhancing the generalizability of findings.

Strengths and limitations

A key strength of this study lies in its innovative crossover design, which enabled rigorous comparison of GenAI and 360° VR simulation across six universities in six countries, drawing on a diverse cohort of nursing students from three different academic years. While the total sample size ($n = 44$) is modest, several factors justify its adequacy. The study achieved a 0% dropout rate, ensuring complete data from all participants. Additionally, the inclusion of students from varied cultural and educational backgrounds enhances the international diversity and relevance of the findings. Moreover, the observed large effect size supports the statistical power of the analyses, as larger effects can be reliably detected with fewer participants, aligning with best practices for sample size justification in resource-constrained or specialized settings. However, limitations must be acknowledged. The small sample, when disaggregated by university, country, and cohort year, restricts the ability to conduct subgroup analyses and may limit the generalizability of the findings. All measures assess perceived competence rather than actual clinical performance through self-report questionnaires. Self-reported competence may not correlate with objective clinical abilities, and future research should incorporate objective assessments alongside self-report tools to provide a comprehensive evaluation of clinical competence. Although participants' English proficiency was not formally assessed, administering the surveys in English is justified, as all participating universities officially adopt English as the medium of instruction. All surveys were administered in English, which served as the common language of instruction across participating institutions. However, because no formal assessment of English proficiency was conducted, there is a possibility that language barriers may have influenced participants' understanding of survey items and the accuracy of their responses, particularly for cultural awareness measures. Future studies should consider formally assessing language proficiency or providing validated translations to ensure linguistic differences do not compromise data validity. The absence of significant changes in the Behavior or Comfort with Interactions and Research Issues subscales may reflect several methodological factors.

Conclusion

The results provide preliminary evidence for the effectiveness of this cutting-edge approach in improving perceived clinical competence, cultural awareness, and AI readiness, with statistically significant improvements observed, particularly within Group B. Both 360° VR simulation and GenAI patient simulation may serve as effective pedagogical strategies for enhancing the clinical outcomes of nursing students. However, GenAI may exhibit a more significant impact, as its initial implementation appeared to engage students more effectively in the learning process, thereby potentially facilitating enhanced educational outcomes. Specifically, when GenAI patient simulation was first introduced, it may have produced superior results; the subsequent use of 360° VR may have further amplified these benefits.

Both groups demonstrated significant improvements across all assessed domains, underscoring the potential efficacy of the intervention in clinical training. The favourable feedback regarding the SET-M suggests that the participating students not only recognised the benefits of these GenAI simulations but also valued the function of GenAI debriefing, which may have enriched their educational experiences and facilitated their professional development.

The integration of GenAI into patient simulations may facilitate dynamic and adaptive learning environments, enabling learners to engage with realistic scenarios that reflect the complexities of actual clinical encounters. Through real-time feedback and GenAI debriefing, this technology may improve diagnostic accuracy and support critical skills such as decision-making and cultural awareness, which are essential elements in today's diverse healthcare landscape.

To fully capitalize on these promising results, future research should focus on identifying the specific components of our GenAI intervention that drive these positive outcomes. By doing so, we may enhance educational strategies in clinical training, ensuring that healthcare professionals are not only technically proficient but also culturally aware and prepared for the challenges of contemporary medical practice. This exploration may be crucial in advancing healthcare education through innovative simulation methodologies.

Acknowledgements
We are grateful for the nursing students participating in the study.

Author contributions
CRediT authorship contribution statement John Fung Tai Chun: Conceptualization, Methodology, Supervision, Writing- original draft and review & editing. Siu Ling Chan: Conceptualization, Methodology, Writing- review & editing. Choi Fung Lam: Methodology, Project administration, Writing- review & editing. Chung Yan Lam: Methodology, Project administration, Writing- review & editing. Christopher Chi Wai Cheng: Methodology, Formal analysis, Writing- review & editing. Man Hin Lai: Methodology, Formal analysis, Project administration, Data Curation, Visualization, Writing- review & editing. Cheuk Chun Joseph Ho: Project administration, Writing- review & editing. Au Siu Lun: Project administration, Writing- review & editing. Mak Lok Yi: Project administration, Writing- review & editing. Sophia Hu: Project administration, Writing- review & editing. Supapak Phetrasuwan: Project administration, Writing- review & editing. Jumpee Granger: Project administration, Writing- review & editing. Jung Min Yoon: Project administration, Writing- review & editing. Gulzar Malik: Project administration, Writing- review & editing. Clara Cabrera Moreno: Project

administration, Writing- review & editing. Patrick Kwok Man Hei: Project administration, Writing- review & editing. Chia-Chin Lin: Project administration, Writing- review & editing.

Funding
The study was funded by the HKU Teaching Development Grant (Grant no 966). The funders did not have a role in study design, data collection, analysis, reporting or the decision to submit for publication.

Data availability
The datasets used and analyzed during the present study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
This study was performed in line with the principles of the Declaration of Helsinki. Ethical approval for this study was obtained from the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (IRB number: UW 24–396). Participants were assured of data confidentiality and voluntariness of participation in and withdrawal from the study, and personal written informed consent was obtained from each of them.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]School of Nursing, Faculty of Medicine, University of Hong Kong, Hong Kong, Hong Kong SAR
[2]School of Health Sciences, Saint Francis University, Hong Kong, Hong Kong SAR
[3]The Chinese University of Hong Kong, Hong Kong, Hong Kong SAR
[4]Hong Kong Baptist University, Hong Kong, Hong Kong SAR
[5]National Yang Ming Chiao Tung University, Taipei, Taiwan
[6]Faculty of Nursing, Mahidol University, Bangkok, Thailand
[7]Ramathibodi School of Nursing, Bangkok, Thailand
[8]Ewha Womans University, Seoul, Republic of Korea
[9]School of Nursing and Midwifery, La Trobe University, Melbourne, Australia
[10]University of Navarra, Pamplona, Spain

Received: 18 March 2025 / Accepted: 26 June 2025
Published online: 17 July 2025

References

1. Laupichler MC, Hadizadeh DR, Wintergerst MW, Emde VD, Paech L, Dick D, E. A., Raupach T. Effect of a flipped classroom course to foster medical students' AI literacy with a focus on medical imaging: A single group pre-and post-test study. BMC Med Educ. 2022;22(1):803.
2. Cross JL, Choma MA, Onofrey JA. (2024). Bias in medical AI: implications for clinical decision-making. PLOS Digit Health. 3(11): e0000651.
3. Agarwal R, Bjarnadottir M, Rhue L, Dugas M, Crowley K, Clark J, Gao G. Addressing algorithmic bias and the perpetuation of health inequities: an AI bias aware framework. Health Policy Technol. 2023;12(1):100702.
4. Parikh RB, Teeple S, Navathe AS. Addressing bias in artificial intelligence in health care. JAMA. 2019;322(24):2377–8.
5. Franco D', Mathew M, Mishra V, Surapaneni KM. Twelve tips for addressing ethical concerns in the implementation of artificial intelligence in medical education. Med Educ Online. 2024;29(1):2330250.
6. Gazzelloni A, Sguanci M, Piredda M, Calandrella C, Tieri G, Piga S, De Marinis MG. Evaluation of the effectiveness of the 360-degree video compared to standard video for nursing students. Preliminary data of a randomized controlled trial. Sci Talks. 2023;6:100201.
7. Pieterse AD, Hierck BP, de Jong PG, Ginn TF, Hamoen EC, Reinders ME. User experiences of medical students with 360-degree virtual reality applications to prepare them for the clerkships. Virtual Reality. 2023;27(2):1381–9.
8. Liaw SY, Tan JZ, Lim S, Zhou W, Yap J, Ratan R, Chua WL. Artificial intelligence in virtual reality simulation for interprofessional communication training: mixed method study. Nurse Educ Today. 2023;122:105718.
9. Alammary MA, Halliday L, Konstantinidis ST. 360-degree virtual reality utilising head-mounted devices in undergraduate nursing and midwifery education: A scoping review. Virtual Worlds. 2023;2(4):396–421.
10. Lee E, Baek G. Development and effects of a virtual reality simulation nursing education program combined with clinical practice based on an information processing model. CIN: Computers Inf Nurs. 2023;41(12):1016–25.
11. Furey P, Town A, Sumera K, Webster CA. (2024). Approaches for integrating generative artificial intelligence in emergency healthcare education within higher education: a scoping review.
12. Seibert K, Domhoff D, Bruch D, Schulte-Althoff M, Fürstenau D, Biessmann F, Wolf-Ostermann K. (2021). Application scenarios for artificial intelligence in nursing care: rapid review. J Med Internet Res. 23(11): e26522.
13. Porter A, Foronda C. Enhancing artificial intelligence literacy in nursing education to combat embedded biases. Nurs Educ Perspect. 2024;45(2):131–2.
14. Lewis AA. Unpacking cultural Bias in AI Language learning tools: an analysis of impacts and strategies for inclusion in diverse educational settings. Int J Res Innov Social Sci. 2025;9(1):1878–92.
15. Currie GM, Hawk KE, Rohren EM. Generative artificial intelligence biases, limitations and risks in nuclear medicine: an argument for appropriate use framework and recommendations. Seminars in nuclear medicine. WB Saunders. 2024.
16. Buchanan C, Howitt ML, Wilson R, Booth RG, Risling T, Bamford M. (2021). Predicted influences of artificial intelligence on nursing education: scoping review. JMIR Nurs. 4(1): e23933.
17. Nazar M, Alam MM, Yafi E, Su'ud MM. A systematic review of human–computer interaction and explainable artificial intelligence in healthcare with artificial intelligence techniques. IEEE Access. 2021;9:153316–48.
18. Mirchi N, Bissonnette V, Yilmaz R, Ledwos N, Winkler-Schwartz A, Del Maestro RF. (2020). The virtual operative assistant: an explainable artificial intelligence tool for simulation-based training in surgery and medicine. PLoS ONE. 15(2): e0229596.
19. Krive J, Isola M, Chang L, Patel T, Anderson M, Sreedhar R. Grounded in reality: artificial intelligence in medical education. JAMIA Open. 2023;6(2):ooad037.
20. Liou SR, Cheng CY. Developing and validating the clinical competence questionnaire: A self-assessment instrument for upcoming baccalaureate nursing graduates. J Nurs Educ Pract. 2014;4(2):56.
21. Gu YH, Xiong L, Bai JB, Hu J, Tan XD. Chinese version of the clinical learning environment comparison survey: assessment of reliability and validity. Nurse Educ Today. 2018;71:121–8. https://doi.org/10.1016/j.nedt.2018.09.026.
22. Rew L, Becker H, Cookston J, Khosropour S, Martinez S. Measuring cultural awareness in nursing students. J Nurs Educ. 2003;42(6):249–57.
23. Yalcinkaya T, Ergin E, Yucel SC. Exploring nursing students' attitudes and readiness for artificial intelligence: A Cross-Sectional study. Teach Learn Nurs. 2024;19(4):e722–8.
24. Eminoğlu A, Çelikkanat Ş. Assessment of the relationship between executive nurses' leadership self-efficacy and medical artificial intelligence readiness. Int J Med Informatics. 2024;184:105386.
25. Karaca O, Çalışkan SA, Demir K. Medical artificial intelligence readiness scale for medical students (MAIRS-MS)–development, validity and reliability study. BMC Med Educ. 2021;21:1–9.
26. Tomesko J, Cohen D, Bridenbaugh J. (2022). Using a virtual flipped classroom model to promote critical thinking in online graduate courses in the United States: a case presentation. Journal of Educational Evaluation for Health Professions. 19.
27. Leighton K, Ravert P, Mudra V, Macintosh C. Updating the simulation effectiveness tool: item modifications and Reevaluation of psychometric properties. Nurs Educ Perspect. 2015;36(5):317–23.
28. Kassis, N., Weber, J. R., Adams, W., Burke, L., Laubham, M. P., Pelka, M.,… Lopez,J. J. (2022). Immersive educational curriculum on intracoronary optical coherence tomography image analysis among naïve readers. BMC Medical Education, 22(1), 719.
29. Zigmont JJ, Kappus LJ, Sudikoff SN. (2011, April). The 3D model of debriefing: defusing, discovering, and deepening. In Seminars in perinatology. 35(2):52-59.
30. Herdman TH, Kamitsuru S, Lopes CT, editors. NANDA-I international nursing diagnoses: definitions & classification, 2024–2026. 13th ed. Thieme. 2024.

31. Lam PPY, Cabalen MB, Peng S, Wong HMG, Botelho MG. (2025). Potentiating the learning outcomes in paediatric dentistry via clinical vicarious learning dialogue videos: A mixed study. Eur J Dent Educ.
32. Saunders WB, Jiang J, Nguyen T. (2007). Linear and generalized linear mixed models and their applications. New York: Springer
33. Zhang P, Kamel Boulos MN. Generative AI in medicine and healthcare: promises, opportunities and challenges. Future Internet. 2023;15(9):286.
34. Rony MKK, Parvin MR, Ferdousi S. (2024). Advancing nursing practice with artificial intelligence: enhancing preparedness for the future. Nurs Open. 11(1).
35. Ličen S, Karnjuš I, Prosen M. Measuring cultural awareness among Slovene nursing student: A cross-sectional study. J Transcult Nurs. 2021;32(1):77–85.

Publisher's note