·数智护理教育专题·

人工智能联合思维引导教学在《健康评估》体格检查教学中的探索研究

张慧[1]　杨梓非[1]　赵益[2]　乔盈志[1]　陈景轩[1]　吕丰[1]　姚秀钰[1]　刘裕文[1]　马伟光[1]
[1]中国医学科学院&北京协和医学院护理学院，北京 100144；[2]南方医科大学护理学院，广州 510515
通信作者：马伟光，Email：zy20061223@126.com

【摘要】 目的　探讨生成式人工智能（artificial intelligence，AI）联合思维引导教学方式在《健康评估》体格检查教学中对护生学习效果和思维能力的影响。**方法**　采用类实验设计，以2个既有教学班（共136人）为研究对象。根据教学资源安排，对其中一个班（69人）实施“生成式AI联合思维引导”教学模式，对另一个班（67人）实施常规教学。研究资料通过《健康评估》理论知识测试卷、体格检查技能操作考核表、体格检查学习成果质量评价表及简化版评判性思维量表进行收集，并采用SPSS 27.0进行统计分析。**结果**　①生成式AI联合思维引导教学模式组的学习成果质量总分高于常规教学组（74.13 ± 9.18 vs. 66.80 ± 11.91，t=2.91，P=0.005），且在学习成果的流程完整性与准确性（14.78 ± 0.54 vs. 13.66 ± 2.21，t=2.92，P=0.006）、临床意义理解（14.08 ± 1.59 vs. 13.03 ± 2.61，t=2.05，P=0.045）、归纳拓展（12.97 ± 1.98 vs. 11.57 ± 2.77，t=2.46，P=0.016）、应用分析维度（11.97 ± 2.81 vs. 10.11 ± 3.59，t=2.43，P=0.018）的得分均高于常规教学组；②生成式AI联合思维引导教学模式组的体格检查技能操作成绩显著高于常规教学组（93.76 ± 3.64 vs. 91.77 ± 6.29，t=2.04，P=0.045），但在理论知识成绩方面与常规教学组差异无统计学意义（84.38 ± 7.87 vs. 82.48 ± 9.07，t=1.30，P=0.194）；③生成式AI联合思维引导教学模式组与常规教学组的评判性思维总分及各维度得分差异均无统计学意义（123.94 ± 11.62 vs. 124.22 ± 12.44，t=−1.53，P=0.130）。**结论**　生成式AI联合思维引导干预在《健康评估》体格检查课程中具有一定的正向促进作用，但AI作为一种教学支持工具，其优化学习过程的教学价值仍有待在更多课程中验证。

【关键词】 人工智能；　思维引导；　健康评估；　护理教育；　教学改革
【中图分类号】 R-05
基金项目：北京协和医学院2025年青年医学教育骨干培训项目（2025mesp008）
DOI: 10.3760/cma.j.cn116021-20251130-02251

Exploration of integrating artificial intelligence with thinking-guided teaching in the physical examination component of the Health Assessment course

Zhang Hui[1], Yang Zifei[1], Zhao Yi[2], Qiao Yingzhi[1], Chen Jingxuan[1], Lü Feng[1], Yao Xiuyu[1], Liu Yuwen[1], Ma Weiguang[1]
[1]School of Nursing, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100144, China; [2]School of Nursing, Southern Medical University, Guangzhou 510515, China
Corresponding author: Ma Weiguang, Email: zy20061223@126.com

【Abstract】 Objective　To explore the effects of a generative artificial intelligence (AI) combined with thinking-guided teaching approach on nursing student learning outcomes and thinking abilities in the physical examination component of the Health Assessment course. **Methods**　A quasi-experimental design was adopted to involve a total of 136 students in two existing classes. Based on the arrangement of teaching resources, one class (n=69) received the generative AI combined with thinking-guided teaching approach, while the other class (n=67) received conventional instruction. Data were collected using a Health Assessment theoretical knowledge test, a physical examination skills assessment checklist, a learning outcome quality evaluation form for physical examination, and a shortened version of the Critical Thinking

Disposition Inventory. Statistical analyses were performed using SPSS 27.0. **Results** ①The total score for learning outcome quality in the generative AI combined with thinking-guided teaching group was significantly higher than that in the conventional teaching group (74.13 ± 9.18 vs. 66.80 ± 11.91, t=2.91, P=0.005). Moreover, the generative AI combined with thinking-guided teaching group scored higher than the conventional teaching group in the dimensions of process integrity and accuracy (14.78 ± 0.54 vs. 13.66 ± 2.21, t=2.92, P=0.006), understanding of clinical significance (14.08 ± 1.59 vs. 13.03 ± 2.61, t=2.05, P=0.045), summarization and extension (12.97 ± 1.98 vs. 11.57 ± 2.77, t=2.46, P=0.016), and application and analysis (11.97 ± 2.81 vs. 10.11 ± 3.59, t=2.43, P=0.018). ②The physical examination skills score of the generative AI combined with thinking-guided teaching group was significantly higher than that of the conventional teaching group (93.76 ± 3.64 vs. 91.77 ± 6.29, t=2.04, P=0.045). However, there was no statistically significant difference in theoretical knowledge scores between the two groups (84.38 ± 7.87 vs. 82.48 ± 9.07, t=1.30, P=0.194). ③No statistically significant differences were observed in the total score of critical thinking or its dimension-specific scores between the two groups (123.94 ± 11.62 vs. 124.22 ± 12.44, t=−1.53, P=0.130). **Conclusions** The generative AI combined with thinking-guided intervention promoted the teaching outcomes of the physical examination component of the Health Assessment course. However, the teaching value of AI as a teaching-assistant tool to optimize the learning process still needs to be validated in more courses.

【**Key words**】 Artificial intelligence; Guided thinking; Health assessment; Nursing education; Teaching reform

Fund program: 2025 Medical Education Scholarship Program (2025mesp008)

DOI: 10.3760/cma.j.cn116021-20251130-02251

近年来，人工智能（artificial intelligence，AI）在高等教育和医学教育中得到广泛应用，其在提升学生学习动机、自主学习能力、知识整合能力及操作技能表现方面的潜在价值已逐渐显现。文献报道，AI辅助教学能够提高学习效率和学习满意度，为医学与护理教育的教学模式创新带来新机遇[1-4]。然而，也有文献报道，学生在学习过程中容易过度依赖AI工具，导致自主学习与独立思考能力下降[5-6]。况且，在现实学习情境中，完全禁止学生使用AI并不切实际[7]。因此，如何以恰当方式将AI融入教学，既发挥其在资源获取和思维支持方面的优势，又避免替代学生必要的思维训练，已成为当前教育研究的关键议题。

健康评估是护理专业学生建立临床思维框架、形成病情识别能力的一门重要基础课程[8]。然而，传统授课模式以讲授和示范为主，教学方式较为单一，导致学生学习兴趣不高、知识掌握零散，临床分析与综合分析能力的培养效果欠佳[9]。护理教育的发展趋势强调培养学生的临床情境综合运用能力、分析推理能力以及决策能力[10]。因此，探索能同时促进知识掌握、技能操作和思维发展的教学策略具有重要意义。基于此，亟须构建一种既能促使学生主动学习，又能在学习过程中给予适度思维支持的教学模式。

本研究拟在《健康评估》课程的体格检查教学中，引入生成式AI与思维引导相结合的教学方法，并采用类实验研究设计，考查该教学模式对护理学生理论学习、技能表现、学习成果质量以及思维能力产生的影响，以期为健康评估课程的教学优化提供科学依据，也为护理教育在AI时代的教学创新提供参考。

1　对象与方法

1.1　研究对象

本研究采用类实验研究设计，以班级为单位开展教学干预。选取2025年9月至11月在北京协和医学院护理学院学习《健康评估》的2024级护理学本科二年级学生（共136名），作为研究对象。根据教学资源与课程安排，对2024级1班（69人）实施“生成式AI联合思维引导”教学模式，对2024级2班（67人）实施常规教学。纳入标准如下：①护理学专业二年级在读学生，初次学习健康评估课程；②学生自愿参加本研究并签署知情同意书；③无严重身体或精神疾病，也无影响学习的重大疾病史。排除标准如下：①曾参与过类似健康评估研究项目；②不愿意参与本研究或未签署知情同意书的护生；

③存在认知或行为障碍，影响数据采集的学生；④正在参与其他AI辅助教学的学生。

1.2 质量控制

为减少选择偏倚和测量偏倚，本研究组建了“生成式AI联合思维引导教学干预团队”，团队成员包括干预实施者、评分者与资料收集者，并对其进行了统一培训与分工。①干预实施者：由第一作者担任，负责设计干预方案、发布思维引导任务以及发放与回收相关资料；②评分者：由6名具有双师型背景的临床教师组成，负责理论知识命题、操作技能考核以及学习报告质量评价。评分时，评分者不知晓学生所在班级的教学模式，以此减少测量偏倚；③资料收集者：由2名未参与干预实施和评分的研究人员担任，负责收集、整理和录入问卷与学习表现数据。

1.3 干预方案

本研究依据课程安排，对2个自然班分别采用不同的教学方式。2个班级均按照相同的教学大纲，完成《健康评估》课程中7个体格检查模块（头颈、淋巴结与皮肤、腹部、呼吸系统、心脏、乳腺、神经系统、脊柱四肢）的理论与技能教学，教学方式存在差异。

1.3.1 干预组：生成式AI联合思维引导教学

干预组在常规教学的基础上，接受将生成式AI支持与思维引导相结合的教学干预。教师依据教学大纲，开展系统的理论讲授与标准化操作示教，并在课堂中明确鼓励学生合理使用AI工具，以此促进学生主动探究、整合知识以及反思。在技能实训环节，教师先进行规范示范，之后学生于模拟情境中开展训练。教师在巡回指导时，鼓励学生借助AI检索资料、优化检查流程、分析临床意义，同时对AI生成内容的准确性展开讨论与判断。学生须在学习过程中记录问题、思考路径以及AI使用情况。在学习报告撰写阶段，学生要依据《健康评估体格检查学习报告质量评分表》，呈现AI在学习过程中的应用情况，具体包括：①检查步骤的验证与优化；②对正常/异常体征临床意义的补充；③思考与问题解决的过程；④AI的作用与局限性；⑤AI与体格检查相结合的创新设想。

课程安排：安排4学时用于学习成果汇报，在此期间，学生展示AI辅助学习的应用经验，教师组织讨论与反思，同时引导学生评价AI在学习过程中的价值。课程结束后，干预组学生须参加统一的理论考试、技能考核以及学习成果质量评定。

1.3.2 对照组：常规教学

对照组采用常规教学模式：教师依照教学大纲，开展系统的理论讲授与标准化操作示教；学生主要借助教材、课堂笔记以及教师提供的参考资料进行复习和训练。在技能实训环节，教师进行规范示范，指导学生在模拟情境中反复练习，及时纠正常见错误，确保技能训练标准化。课程中同样安排4学时的学习汇报，学生整理并展示所学知识与技能经验，教师组织讨论，促进互学互鉴。

在整个教学过程中，未明确支持或要求学生使用AI工具，学生主要依靠传统学习资源来完成课程学习。课程结束后，对照组学生参加了统一的理论考试与技能操作考核。

1.4 评价方法

1.4.1 研究工具

1.4.1.1 理论知识测试卷

理论知识测试依据《健康评估》课程教学大纲及核心知识点，由研究团队构建题库。各系统单元题库包含单项选择题与多项选择题。正式考核时，通过在线平台发放试题，系统从题库中自动抽取25道（单项选择题20道、多项选择题5道），每题设4个备选答案，每题4分，总分100，答题时限为60 min。系统会自动对题目顺序和选项顺序进行乱序处理，以降低记忆干扰效应。得分越高，表明学生对体格检查理论知识的掌握程度越高。

1.4.1.2 技能操作考核

使用研究团队统一制定的《体格检查技能操作考核评分表》评价操作能力。考核项目涵盖头颈淋巴结与皮肤、腹部、呼吸系统、心脏、乳腺、神经系统、脊柱四肢7个模块，总分100，得分越高，技能操作水平越高。

1.4.1.3 评判性思维量表

采用经国内修订的简化版评判性思维倾向量表（Chinese critical thinking disposition inventory，CTDI-CV）进行测量[11]。该量表共设28个条目，涵盖自信心、求知欲、系统化能力、认知成熟度、分析能力和寻求真相6个维度。量表采用1~6分Likert评分法（1分=非常不同意，6分=非常同意），得分越高，表明评判性思维倾向越强。既往研究表明，该量表在医学生群体中总量表的Cronbach’s α 系数为0.865，各维度系数为0.681~0.917[12]。

1.4.1.4 学习成果质量评价表

采用研究团队自编的《体格检查学习成果质量

评价表》来评价学生的学习产出。该量表包含操作流程质量、对临床意义的理解、问题与思考、归纳拓展及应用分析5个维度，共19个条目。其中，第12、18、19条目仅适用于干预组，在统计分析时不计入总分。扣除这些条目后，总分范围为17~85分，得分越高，表明学习成果质量越高。学习成果由6名经过统一培训的教师采用盲法独立评分，最终得分取平均值。评分者间一致性系数为0.908，量表内容效度指数为0.936，这表明该评价工具具有良好的一致性和内容效度。

1.4.2　资料收集

本研究所有资料均通过在线平台，由调查对象以自填方式完成收集。资料收集工作由课程负责人在预设时间节点统一发放问卷，并由经过培训的研究助理进行现场督导，以确保数据的完整性与准确性。资料收集者均未参与教学干预或评分工作，保证了数据收集过程的独立性。

根据研究设计，资料收集分为3个时间节点（图1）。T1（课程开始前）：收集2组学生的基线资料，包括性别、年龄、既往学业成绩等一般人口学信息，并进行评判性思维倾向测评，以此掌握2组学生的初始水平。T2（完成6次课堂教学后）：开展学习成果汇报，汇报结束后即时收集学习成果质量评价结果以及第2次评判性思维量表数据，用以反映学生在学习过程中的阶段性表现和思维变化。T3（课程结束时）：组织统一的技能操作考核，操作考核后收集理论知识测试成绩，获取最终的知识掌握和技能表现数据。

1.5　统计学方法

使用SPSS 27.0软件进行数据处理与分析。计数资料以频数和百分率表示，组间比较采用卡方检验。计量资料首先进行正态性检验（Shapiro-Wilk）。符合正态分布的以均数 ± 标准差表示，组间比较采用独立样本 t 检验；不符合正态分布的以中位数和四分位数表示，组间比较采用Mann-Whitney U 非参数检验（以 Z 值报告）。本研究为类实验设计，组间比较的统计推断结果仅作为探索性参考。所有检验均为双侧检验，检验水准 α=0.05。

2　结果

2.1　2组一般资料比较

本研究最终纳入136名护生。依据课程安排，将2024级1班（69人）设为干预组，2024级2班（67人）设为对照组。学生年龄为19（19，20）岁，其中女生95人（69.85%），男生41人（30.15%）；生源地为城市的31人（22.79%），农村生源的105人（77.21%）。入校成绩为599（586，611）分，担任班干部的有35人（25.74%）。对2组学生的一般资料进行组间比较，结果显示：年龄、性别、生源地、班干部经历及入校成绩等基线特征在2组间差异均无统计学意义（P>0.05），即2组基线水平相似（表1）。

2.2　干预组和对照组护生评判性思维基线比较

对干预组和对照组的评判性思维总分及各维度得分进行基线比较，结果显示：2组在总分及6个维度上的差异均无统计学意义（P>0.05），即2组在评判性思维的初始水平上相似（表2）。

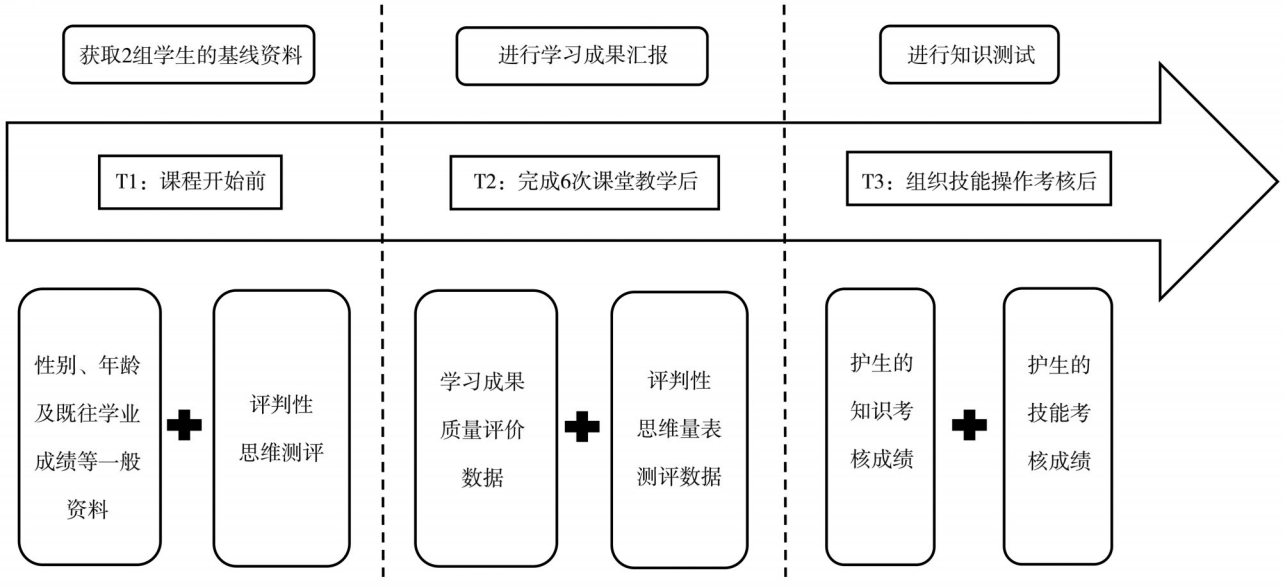


图1　资料收集的时间节点

表 1　2 组护生一般资料比较

项目	总数（n=136）	干预（n=69）	对照组（n=67）	χ^2/Z值	P值
年龄[岁，$M(Q_1,Q_3)$]	19（19，19）	19（19，19）	19（19，19）	−0.55[a]	0.581
性别[例（%）]				0.09[b]	0.765
男	41（30.15）	20（28.99）	21（31.34）		
女	95（69.85）	49（71.01）	46（68.66）		
生源地[例（%）]				1.79[b]	0.181
城市	31（22.79）	19（27.54）	12（17.91）		
农村	105（77.21）	50（72.46）	55（82.09）		
入校时成绩[分，$M(Q_1,Q_3)$]	599（584，609）	598（592，610）	599（582，609）	−0.95[a]	0.341
是否担任班干部[例（%）]				1.17[b]	0.279
是	35（25.74）	15（21.74）	20（29.85）		
否	101（74.26）	54（78.26）	47（70.15）		

注：a，Z值；b，χ^2值。

表 2　干预前 2 组护生评判性思维总分和各维度得分比较

项目	干预（n=69）	对照组（n=67）	Z/t值	P值
评判性思维[分，$M(Q_1,Q_3)$]	122.00（114.50，137.00）	125.00（117.00，135.00）	−0.54[a]	0.588
自信心（分，$\bar{x}\pm s$）	30.78 ± 4.74	31.48 ± 4.20	−0.90[b]	0.367
求知欲（分，$\bar{x}\pm s$）	22.72 ± 3.51	23.48 ± 3.31	−1.29[b]	0.201
系统化能力（分，$\bar{x}\pm s$）	21.93 ± 3.31	21.90 ± 3.52	0.06[b]	0.956
认知成熟度（分，$\bar{x}\pm s$）	22.33 ± 3.35	21.99 ± 4.03	0.55[b]	0.585
分析能力[分，$M(Q_1,Q_3)$]	14.00（13.00，16.00）	15.00（14.00，16.00）	−1.59[a]	0.112
寻找真相（分，$\bar{x}\pm s$）	12.70 ± 2.15	12.12 ± 2.62	1.40[b]	0.163

注：a，Z值；b，t值。

2.3　干预后 2 组护生体格检查理论知识成绩、操作技能成绩及学习成果质量比较

干预结束后，对 2 组护生的理论知识成绩、操作技能成绩以及学习成果质量得分进行比较。结果显示，干预组学习成果质量总分高于对照组（干预组 86.11 ± 11.35，对照组 76.34 ± 14.27，P=0.003）。在学习成果的多个维度，如流程完整性与准确性、对临床意义的理解、问题与思考、归纳拓展及应用分析等，干预组得分均高于对照组（P<0.05）。此外，干预组的操作技能成绩也高于对照组，差异具有统计学意义（表 3）。

表 3　干预后 2 组护生学习成效的比较（分，$\bar{x}\pm s$）

项目	干预组（n=69）	对照组（n=67）	t值	P值
理论知识	84.38 ± 7.87	82.48 ± 9.07	1.30	0.194
操作技能	93.76 ± 3.64	91.77 ± 6.29	2.04	0.045
学习成果质量	74.13 ± 9.18	66.80 ± 11.91	2.91	0.005
操作流程质量	14.78 ± 0.54	13.66 ± 2.21	2.92	0.006
临床意义的理解	14.08 ± 1.59	13.03 ± 2.61	2.05	0.045
问题与思考	20.33 ± 4.69	18.43 ± 3.79	1.88	0.064
归纳拓展	12.97 ± 1.98	11.57 ± 2.77	2.46	0.016
应用分析	11.97 ± 2.81	10.11 ± 3.59	2.43	0.018

2.4　干预后干预组和对照组护生评判性思维得分比较

干预结束后，对 2 组护生的评判性思维总分及各维度得分进行比较。结果显示，2 组在评判性思维总分及 6 个维度上的差异均无统计学意义（表 4）。

表 4　干预后 2 组护生评判性思维总分和各维度得分比较（分，$\bar{x}\pm s$）

项目	干预组（n=69）	对照组（n=67）	t值	P值
评判性思维	123.94 ± 11.62	124.22 ± 12.44	−1.53	0.130
自信心	30.07 ± 4.93	31.31 ± 4.54	−0.29	0.769
求知欲	22.78 ± 3.43	22.96 ± 3.42	0.32	0.749
系统化能力	21.62 ± 3.33	21.43 ± 3.60	1.31	0.194
认知成熟度	22.71 ± 3.71	21.85 ± 3.96	−0.63	0.529
分析能力	14.06 ± 2.24	14.30 ± 2.21	0.76	0.448
寻找真相	12.70 ± 2.54	12.37 ± 2.40	−0.14	0.892

2.5　干预组护生干预前后评判性思维得分比较

干预结束后，对干预组护生的评判性思维总分及各维度得分展开比较。结果显示，干预前后，该组护生评判性思维总分及 6 个维度的得分差异均无统计学意义（表 5）。

表5　干预组护生干预前后评判性思维总分和各维度得分比较(分,$\bar{x} \pm s$)

项目	干预前(n=69)	干预后(n=69)	t值	P值
评判性思维	124.71 ± 12.12	123.94 ± 11.62	0.77	0.692
自信心	30.78 ± 4.74	30.07 ± 4.93	0.71	0.338
求知欲	22.72 ± 3.51	22.78 ± 3.43	–0.06	0.916
系统化能力	21.93 ± 3.31	21.62 ± 3.33	0.30	0.587
认知成熟度	22.33 ± 3.35	22.71 ± 3.71	–0.38	0.514
分析能力	14.25 ± 2.26	14.06 ± 2.24	0.19	0.591
寻找真相	12.70 ± 2.15	12.70 ± 2.54	0.00	1.000

3　讨论

3.1　AI辅助可能有助于提升操作技能表现,其优势或与即时反馈和情境模拟相关

本研究结果显示,干预组学生在体格检查操作技能和学习成果质量方面的得分更高,这提示生成式AI联合思维引导教学可能有助于提升学生在操作性学习任务中的表现。既往研究表明,将AI嵌入情境学习、反思讨论或临床模拟活动中,能够提高学习者的投入度和任务准备度[13]。本研究的数据与这些研究结果一致,说明AI在支持技能学习方面可能具有一定潜力。这种潜在的促进作用,或许与AI提供的学习支持机制有关,例如即时反馈、可重复操作的模拟情境,以及能让学生在训练中即时修正并验证结果、整合多元资料的学习体验等[14]。此外,AI能提供高度个性化的反馈和多样化的学习方式,使学生可根据自身需求调整学习节奏,进一步优化技能学习过程。这种潜在促进作用可能不仅体现在操作性任务上,还能间接提升学生在临床情境中的思维建构与决策能力。不过,由于本研究采用类实验设计,该结论仍需在更严格的实验条件下进一步验证。

3.2　AI可能更有利于支持理解性学习,而对记忆性学习影响有限

研究显示,干预组学生的理论知识成绩差异无统计学意义,这与部分国外研究结果一致,即AI在知识学习方面虽呈现积极趋势,但在提升标准化考试成绩上效果有限[14]。这或许是因为AI在教学中主要提供解释性、分析性或结构化信息,有助于促进学生对内容的理解与应用,但对于以机械记忆和固定答案为主的理论测试,其直接作用相对有限[15]。此外,学生在使用AI时,往往更关注知识间的逻辑关系与临床意义,而非重复记忆知识点。国内研究指出,《健康评估》课程教学改革应更注重学习过程与思维建构,而非仅将理论成绩作为教学效果的唯一衡量指标[16]。本研究进一步验证了这一观点,表明在理论教学中整合AI时,教学活动设计应更注重培养学生理解、分析和应用能力,避免仅将AI当作信息提供工具,而应通过有效教学策略推动学生对知识的深入理解与思维迁移。这一发现为今后AI在护理教育中的进一步应用指明了方向,即需在课程设计中融入更复杂的认知任务和跨情境的推理活动,才能最大化发挥AI的潜力。

3.3　AI对高阶思维的促进可能需要更复杂的情境与更长周期的教学设计

本研究结果显示,干预组护生评判性思维总分及各维度得分在干预前后差异无统计学意义;干预后,干预组与对照组间评判性思维水平差异也无统计学意义。这一结果提示,在本研究所设定的教学周期内,评判性思维水平的变化幅度相对有限,短期内难以在量化测评工具上呈现出稳定且显著的提升。从能力发展特征来看,既往研究表明,评判性思维的形成通常需要经历多情境的冲突体验、跨情境推理、持续反思,且需要较长的学习周期[17]。相比之下,本研究中的AI联合思维引导教学主要嵌入单门课程的课堂教学与技能训练情境中,干预周期较短,且尚未系统引入复杂临床决策、角色冲突模拟或跨情境推演等高阶认知任务,这可能在一定程度上限制了其对评判性思维发展的促进作用。此外,本研究未对学生使用AI的频率、方式及其在学习过程中的参与深度进行过程性量化分析,因此尚无法从数据层面进一步探讨AI使用行为与评判性思维变化之间的关联,这在一定程度上限制了对作用机制的深入解释。研究报道,若AI的使用缺乏明确的思维引导与任务约束,其高度便利性在部分学习情境中可能降低学生主动推理与深度思考的投入,从而对高阶思维发展产生中性甚至削弱效应[18–19]。由此可见,未来研究有必要在延长干预周期的基础上,结合多情境教学设计、复杂临床案例推演,并对AI使用行为进行过程性记录,以更全面地评估AI在促进高阶思维发展中的潜在价值。

3.4　本研究的局限性

本研究存在以下局限性:①本研究采用类实验设计,研究结果主要呈现关联性趋势,无法据此进行因果推断。未来可采用更多整群随机对照试验或多中心多层分析方法,增强推断的严谨性。②本研究为单中心研究,样本规模有限,外部效度较低。未来需在不同院校、更大样本中开展研究,提升研

究结论的普适性。③本研究使用的学习成果质量评价表为自编评价工具，虽已验证其内容效度，但因样本量有限，尚未系统检验其结构效度。未来需进一步验证该量表的结构效度，提高该评价表的应用可靠性。

4　结语

本研究结果提示，在《健康评估》体格检查课程情境中，生成式 AI 支持的思维引导对操作性技能和学习成果质量有一定促进作用，在理解性知识学习方面也表现出积极趋势，但短期内未能使理论考试成绩显著提升，其对评判性思维等高阶认知能力的影响也较为有限。鉴于本研究采用类实验设计，研究对象集中于同一年级学生，且教学情境主要限于单门课程内，相关结果仅能反映特定课程条件下的关联性趋势。鉴于此，在以技能训练与情境理解为核心的课程中，应合理整合生成式 AI，并辅以明确的思维引导。AI 作为教学支持工具优化学习过程的教学价值，仍需在更大样本、多课程研究中进一步验证。

利益冲突　所有作者声明不存在利益冲突

作者贡献声明　张慧：提出研究构想与研究设计；张慧、杨梓非、赵益：论文撰写与修订；张慧、乔盈志、陈景轩、吕丰：研究实施；姚秀钰、刘裕文：数据收集与统计分析；马伟光：研究总体指导、质量控制、论文修改与定稿

参考文献

[1] Sharma M, Sharma S. A holistic approach to remote patient monitoring, fueled by ChatGPT and metaverse technology: the future of nursing education [J]. Nurse Educ Today, 2023, 131: 105972. DOI: 10.1016/j.nedt.2023.105972.

[2] Lin MY, Huang MZ, Lai PC. Effect of virtual reality training on clinical skills of nursing students: a systematic review and meta-analysis of randomized controlled trials [J]. Nurse Educ Pract, 2024, 81: 104182. DOI: 10.1016/j.nepr.2024.104182.

[3] Chen PJ, Liou WK. ChatGPT-driven interactive virtual reality communication simulation in obstetric nursing: a mixed-methods study [J]. Nurse Educ Pract, 2025, 85: 104383. DOI: 10.1016/j.nepr.2025.104383.

[4] Srinivasan M, Venugopal A, Venkatesan L, et al. Navigating the pedagogical landscape: exploring the implications of AI and chatbots in nursing education [J]. JMIR Nurs, 2024, 7: e52105. DOI: 10.2196/52105.

[5] Alenazi L, Alhalal E. Factors affecting artificial intelligence usage intention among nursing students: unified theory of acceptance and use of technology [J]. Nurse Educ Today, 2025, 152: 106780. DOI: 10.1016/j.nedt.2025.106780.

[6] Tam W, Huynh T, Tang A, et al. Nursing education in the age of artificial intelligence powered chatbots (AI-chatbots): are we ready yet [J]. Nurse Educ Today, 2023, 129: 105917. DOI: 10.1016/j.nedt.2023.105917.

[7] Gordon M, Daniel M, Ajiboye A, et al. A scoping review of artificial intelligence in medical education: BEME Guide No. 84 [J]. Med Teach, 2024, 46(4): 446-470. DOI: 10.1080/0142159X.2024.2314198.

[8] Wang X, Dong L, Lyu W, et al. Teaching health assessment symptomatology using a flipped classroom combined with scenario simulation [J]. J Nurs Educ, 2020, 59(8): 448-452. DOI: 10.3928/01484834-20200723-05.

[9] 陈萌. 激发与维持学习动机教学策略在健康评估课程中的应用效果评价[J]. 中华护理教育, 2019, 16(3): 200-203.

[10] Garrett BM. Changing the game; some thoughts on future healthcare demands, technology, nursing and interprofessional education [J]. Nurse Educ Pract, 2012, 12(4): 179-181. DOI: 10.1016/j.nepr.2012.03.006.

[11] 于斌斌, 于良芝. 面向LIS研究与实践的批判思维倾向量表：对CTDI-CV进行简化的可行性检验[J]. 图书与情报, 2020(4): 21-27.

[12] 杨光媚, 韦慧燕, 万乐平, 等. 简化版批判性思维量表在医学生中的信效度分析[J]. 郑州大学学报(医学版), 2022, 57(3): 330-335. DOI: 10.13705/j.issn.1671-6825.2021.07.104.

[13] Cucci F, Marasciulo D, Romani M, et al. The contribution of artificial intelligence in nursing education: a scoping review of the literature [J]. Nurs Rep, 2025, 15(8): 283. DOI: 10.3390/nursrep15080283.

[14] Jembu JP, Lee S. The effectiveness of artificial intelligence (AI) in learning outcomes of nursing students: a meta-analysis of randomized controlled trials (RCTs) [J]. Steme, 2025, 5(4): 546-563. DOI: 10.3934/steme.2025026.

[15] Khlaif ZN, Salameh N, Ajouz M, et al. Using generative AI in nursing education: students' perceptions [J]. BMC Med Educ, 2025, 25(1): 926. DOI: 10.1186/s12909-025-07416-z.

[16] 王梦瑶, 何悦, 侯峣, 等. 基于OSCE考核路径的《健康评估》课程教学质量聚类分析[J]. 吉林医学, 2025, 46(10): 2605-2609.

[17] 张艳青, 李冬莉, 孔静, 等. 护理本科生评判性思维能力研究进展[J]. 临床合理用药杂志, 2014, 7(20): 178-179. DOI: 10.15887/j.cnki.13-1389/r.2014.20.008.

[18] Preiksaitis C, Rose C. Opportunities, challenges, and future directions of generative artificial intelligence in medical education: scoping review [J]. JMIR Med Educ, 2023, 9: e48785. DOI: 10.2196/48785.

[19] Allen KA, Costner-Lark A, Serratt T, et al. Artificial intelligence in advanced practice nursing education: opportunities and challenges [J]. J Am Assoc Nurse Pract, 2026, 38(2): 102-107. DOI: 10.1097/JXX.0000000000001169.

（收稿日期：2025-11-30）
（本文编辑：兰明娟）