

大语言模型在护理学领域应用的范围综述

马应卓¹, 王俊¹, 刘彤¹, 蒋婧¹, 吴洁¹, 赵梓寒¹, 赵庆华¹, 肖明朝²

摘要: **目的** 对大语言模型在护理学领域的应用研究进行范围综述,为相关研究提供参考。**方法** 遵循范围综述方法,系统检索中英文数据库中相关大语言模型在护理学领域应用的相关研究。检索时限为建库至 2024 年 5 月,由 2 名研究人员独立筛选提取资料,并对纳入文献进行总结和分析。**结果** 共纳入 33 篇文献,大语言模型在护理学领域的应用主要集中于护理教育(48.49%)、临床护理(42.42%),护理科研(3.03%)及综合应用(6.06%);其应用形式主要包括护理专业考试或护理基础知识测试、护理教学写作、临床模拟场景构建、制订护理计划、协助患者病情管理、为患者提供专业知识和信息支持等。大语言模型已经在护理教育和临床护理的患者管理方面展现出了显著的应用效果,但在知识应用、回答临床问题、进行专业判断或决策时仍有缺陷,不足以取代人工。**结论** 大语言模型在我国护理学领域的研究处于起步阶段,未来的研究应当积极整合前沿技术,以促进其在该领域的深入应用,同时,应加快制定和推行标准化的使用准则和指南,确保技术应用的安全性和有效性,切实推动护理学领域的创新与发展。**关键词:** 大语言模型; 人工智能; ChatGPT; 护理教育; 护理科研; 临床护理; 范围综述
中图分类号: R47; TP18 **DOI:** 10.3870/j.issn.1001-4152.2024.19.124

A scoping review of the applications of large language models in nursing science

Ma Yingzhuo, Wang Jun, Liu Tong, Jiang Jing, Wu Jie, Zhao Zihan, Zhao Qinghua, Xiao Mingchao. Department of Nursing, The First Affiliated Hospital of Chongqing Medical University, Chongqing 400016, China

Abstract: **Objective** To conducted a scoping review to summarize and describe the application of large language models in nursing, and to provide a reference for relevant research. **Methods** Following the scoping review methodology, relevant studies were identified through searches in Chinese and English databases from the construction to May 2024. Two researchers independently screened and extracted data, summarized and analyzed the included literatures. **Results** A total of 33 studies were included. The applications of large language models in nursing primarily focused on nursing education (48.49%), clinical care (42.42%), nursing research (3.03%) and comprehensive applications(6.06%), the forms of its application mainly included nursing professional exams or tests of basic knowledge of nursing, nursing didactic writing, construction of clinical simulation scena-rios, development of nursing care plans, assisting in the patient management, and providing professional knowledge and information support for patients. While large language models had shown significant effectiveness in nursing education and patient management in clinical nursing, they still had limitations in knowledge application, answering clinical questions, and making professional judgments or decisions, and were not yet sufficient to replace human involvement. **Conclusion** Research on large language models in the nursing field in China is still in its early stages. Future studies should focus on actively integrating cutting-edge technologies to facilitate their deeper application in this area. Furthermore, there is a pressing need to develop and implement standardized guidelines to ensure the safety and effectiveness of these technologies, thereby driving innovation and advancement in nursing.
Keywords: large language models; artificial intelligence; ChatGPT; nursing education; nursing research; clinical nursing; scoping review

人工智能(Artificial Intelligence, AI)技术在护理学领域已经发展近 40 年,其在改善临床护理质量、赋能护理教育发展、辅助护理科学研究等方面展现了巨大潜力^[1]。中国知网中首次提及人工智能和护理学的交叉应用可以追溯到 1985 年,该研究采用人工智能技术研发了责任制护理软件系统^[2]。近年来,人工智能技术的快速发展促进了大语言模型(Large

Language Models, LLM)的产生和出现。大语言模型是指包含数十到万亿甚至更多参数的 Transform 语言模型,这些模型主要使用深度学习技术,经过大规模的数据训练,能够执行自然语言处理和人机交互等复杂的任务^[3]。近期较为著名的是 2022 年 11 月由美国 Open AI 公司发行的大语言模型:ChatGPT,一经推出便引起了医疗及护理领域的广泛关注^[4],之后 OpenAI 公司又于 2024 年 5 月正式发布了最新的旗舰模型——ChatGPT 4o,它提供了更强大的自然语言理解、上下文感知和集成功能^[5]。与此同时,国外的谷歌、微软等科技公司研发了 Bard, Bing-Chat (2023 年 11 月更名为 Copilot),国内百度、科大讯飞及清华大学、复旦大学等研究团队积极对接新技术和新理念研发了本土化的大语言模型,如百度的文心一言。目前,大语言模型在医疗领域研究较为广泛,但

作者单位:重庆医科大学附属第一医院 1. 护理部 2. 泌尿外科 (重庆,400016)
马应卓:男,硕士,护士,yingzhuo1999@126.com
通信作者:肖明朝,xmz.2004@163.com
科研项目:重庆医科大学 2019 年校级教育教学研究项目 (JY190316);重庆市高等教育教学改革研究项目重点项目 (222069);重庆市教委十四五重点学科支持项目(20240101)
收稿:2024-05-24;修回:2024-07-20

护理学领域尚处于探索阶段。因此,本研究基于 Arksey 等^[6]的范围综述方法学框架,对大语言模型在护理学领域中的相关应用研究进行范围综述,旨在全面了解大语言模型在护理学领域的研究现状和发展趋势,为我国护理学者开展相关研究提供基础依据。

1 资料与方法

1.1 确定研究问题 具体研究问题为:①大语言模型在护理学领域的具体应用场景是什么? ②大语言模型在护理学领域的潜在应用价值和局限性包括什么? ③针对大语言模型在护理学领域应用的策略及意见有哪些?

1.2 文献纳入与排除标准 纳入标准:①内容涉及大语言模型在护理学领域的应用;②可获取全文;③中英文文献。排除标准:①未明确指出人工智能产品是否基于大语言模型;②重复发表、信息及数据不完整;③无法获得全文;④社评、评论、简讯、观点类等非学术性研究类文章。

1.3 文献检索 以大语言模型,生成式人工智能,自然语言处理,ChatGPT, chatgpt, GPT, gpt, Google Bard, Bard, Bing-Chat, Copilot, LaMDA, LLaMA, MT-NLG, BatGPT, Moss, Sora, Kimi, 文心一言, 通义千问, 讯飞星火, 智谱清言, 天工, 腾讯元宝, 紫东太初; 护理学, 护理, 护理教育, 护理科研, 临床护理为中文检索词。以 natural language processing, large language model *, generative artificial intelligence, LLM,

Generative Pre-trained Transformer, natural language generat *, ChatGPT *, chatgpt, GPT *, gpt, Google Bard, Bard, Bing-Chat, Copilot, LaMDA, LLaMA, MT-NLG, BatGPT, Moss, Sora, Kimi, Ernie Bot, Qwen *, iFLYTEK Spark, ChatGLM *, Skywork, Tencent Yuanbao, ZiDongTaiChu; Nurs *, registered nurses, RN, advanced practice nurse practitioner, ARNP, clinical nursing, nursing education, nursing research 为英文检索词,在中国知网、万方数据、维普网、PubMed、Web of Science、The Cochrane Library、CINAHL、Embase 中英文数据库中进行检索。检索时限为建库至 2024 年 5 月。此外,补充检索纳入文献的参考文献,并在相关网站上追踪检索最新发表的文献。

1.4 文献筛选及资料提取 将检索出的文献导入 EndNote 软件去重,由 2 名参与过循证护理培训的研究者严格按照纳入与排除标准进行文献的筛选和内容的提取,结果有争议时由课题组专家进行判定。文献提取的内容包括:作者、国家、研究设计、研究内容、研究领域、研究结果。

2 结果

2.1 文献筛选结果 经过初步检索获得相关文献共 4 790 篇。去重后获得文献 3 933 篇;阅读文献题目、摘要后剔除 3 622 篇,阅读全文后排除与研究主题不相关文献 98 篇、无法获取全文 20 篇、内容重复或相似 63 篇、不符合纳入标准 97 篇,最终纳入 33 篇。

2.2 纳入文献基本特征 见表 1。

表 1 纳入文献基本特征

作者	国家	研究方法	研究领域	工具	研究内容	研究结果
Kaneda 等 ^[7]	日本	描述性研究	临床护理	GPT-4、GPT-3.5	护理知识测验	GPT-4 准确率较 GPT-3.5 更高,并符合日本国家护理考试通过标准
Taira 等 ^[8]	日本	描述性研究	临床护理	GPT-3.5	护理知识测验	GPT-3.5 符合 2019 年日本国家护士考试的通过标准,接近通过 2020—2023 年的考试标准
Chan 等 ^[9]	中国	其他	护理教育	ChatGPT *	学生使用 ChatGPT 解决护理案例中的具体问题	学生表达了 ChatGPT 的获益感和风险性,并强调了批判性思维和专业判断能力在护理专业的重要性
Parker 等 ^[10]	美国	描述性研究	护理科研	GPT-3	GPT-3 作为学生写作的评价工具的有效性	GPT-3 对护理学生的写作提供了详细且类似于人类的反馈,提出具体的改进建议,但评分比人工更严格
Cox 等 ^[11]	美国	描述性研究	护理教育	GPT-3.5	比较人工智能和护理教师回答护士执照考试的水平	GPT-3.5 生成的美国注册护士执照考试题在清晰度、语法、临床相关性和术语复杂度上与护理教育者相当,但存在一定差异
Hsu ^[12]	中国	实验性研究	护理教育	ChatGPT *、Term-bot	评估 ChatGPT 和 Term-bot 在改善护生学习医学术语方面的效果	ChatGPT 和 Termbot 等工具可提升护生学习效果,成为护生的学习工具
Quattrini 等 ^[13]	美国	未提及	护理教育	ChatGPT *	评价 ChatGPT 在护理教学课堂的应用效果	ChatGPT 在护理课程中有效提升了学生的临床判断能力,特别是在高阶认知评估中,但仍需考虑其风险性
Branum 等 ^[14]	美国	描述性研究	综合	ChatGPT *	评价 ChatGPT 用于回答复杂临床问题的效果	ChatGPT 当前版本不适合用于回答临床问题,特别是在需要证据支持的情况下,未能取代人工进行证据检索和研究
夏浩然等 ^[15]	中国	回顾性研究	临床护理	ChatGPT *	ChatGPT 平台用于瓣膜病患者自我监测随访系统	搭载 ChatGPT 的大数据随访平台可提高护理工作效率,协助患者病情管理
Huang ^[16]	中国	描述性研究	护理教育	GPT-3.5	护理知识测验	GPT-3.5 可以迅速对题目做出判断,并通过了考试,但会产生误导性或不准确的解释
刘亚玲等 ^[17]	中国	描述性研究	临床护理	ChatGPT *、新必应	消化内镜癌症知识测验	新必应回答问题准确性更高,但在内镜不良事件问题和食管癌筛查方法问题回答方面表现不佳

续表 1 纳入文献基本特征

作者	国家	研究方法	研究领域	工具	研究内容	研究结果
Wang 等 ^[18]	中国	描述性研究	临床护理	GPT-3.5、GPT-4	气管造口护理相关知识测验	GPT-4 在各类题目中表现都更佳,平均超过 50% 的准确率
Luo 等 ^[19]	中国	描述性研究	护理教育	ChatGPT [*]	中国护生对人工智能工具 ChatGPT 的认知和需求	中国部分护生对 ChatGPT 的熟悉程度较低,使用经验不足
谭晓文等 ^[20]	中国	描述性研究	临床护理	文心一言、智谱清言、讯飞星火、通义千问	对前列腺癌围术期护理与健康教育相关问题咨询响应	通义千问等大型语言模型可在前列腺癌围术期护理中为患者提供专业知识和信息支持
Karacan ^[21]	土耳其	描述性研究	临床护理	LaMDA、GPT-4	评价其在制订产后出血护理计划的效果	与 LaMDA 和 GPT-4 相比,Med-PaLM 在制订护理计划方面表现出更高的潜力,但并不适用于真实患者
Chang 等 ^[22]	中国	描述性研究	护理教育	ChatGPT [*]	将 ChatGPT 系统融入到护理与健康教育设计课程的教学活动	ChatGPT 系统有效提高了护生的分析判断能力、解决问题的能力和学习乐趣
Alanzi ^[23]	沙特阿拉伯	质性研究	临床护理	ChatGPT [*]	ChatGPT 作为医疗护理远程咨询工具	共凝练 12 个积极主题和 6 个消极主题,为类 ChatGPT 等大语言模型融入临床实践提供具体参考意见
Zong 等 ^[24]	中国	描述性研究	护理教育	GPT-3.5-Turbo	护理知识测验	GPT-3.5-Turbo 未能通过 2017—2021 中国执业护士考试
Saban 等 ^[25]	以色列	描述性研究	临床护理	GPT-3.5	运用于辅助临床护士决策	GPT-3.5 在临床和护理决策中表现出优柔寡断和过度分类
Su 等 ^[26]	中国	描述性研究	护理教育	GPT-4	护理知识测验	GPT-4 通过护理考试,整体准确率 80.75%,且在不同类型题目中表现具有差异性
Vaughn 等 ^[27]	美国	未提及	护理教育	ChatGPT [*]	运用于创建准确且真实的模拟场景,协助教师工作	ChatGPT 有潜力帮助模拟护理教育者开发临床模拟场景,但需要专家的审核和修改以确保场景的真实性和全面性
Zaboli 等 ^[28]	意大利	描述性研究	临床护理	GPT-3.5	运用于患者分诊	GPT-3.5 不足以替代分诊护士在急诊患者优先处理中的专业判断
Woodley ^[29]	美国	未提及	护理教育	ChatGPT [*]	运用于护理教育课程	ChatGPT 融入护理研究生课程能够改善教学效果,促进临床技能的培养
Sblendorio 等 ^[30]	意大利	未提及	临床护理	大语言模型	评估各种大语言模型应用于临床护理的可行性	Gemini Advanced、Claude 3 和 GPT-4 被分类为“推荐”。LLAMA 3-70B 和 GPT-3.5 被归类为“高度谨慎使用”
Yahagi 等 ^[31]	美国	实验性研究	临床护理	GPT-3.5	用于缓解全身麻醉成人手术患者术前焦虑	GPT-3.5 显著降低了患者术前焦虑,然而焦虑评分没有整体差异,此外强调了改进 GPT 算法和知识库以提高性能和满意度的必要性
Dubin 等 ^[32]	美国	实验性研究	临床护理	GPT-3.5	用于回答全髋关节置换术后患者常见问题	大部分患者对 GPT-3.5 的相关回答更加满意
Levin 等 ^[33]	以色列	描述性研究	临床护理	GPT-4、Claude-2.0	用于新生儿护理场景下的临床推理	Claude-2.0 在新生儿临床护理推理能力方面优于 GPT-4,但两种模型都在诊断精度、治疗特异性等存在局限性
穆兰等 ^[34]	中国	描述性研究	护理教育	ChatGPT [*] 、文心一言	护理知识测验	大语言模型在护理基础知识测试中表现优异,但在知识应用和独立思考方面存在缺陷
Dagci 等 ^[35]	土耳其	描述性研究	综合	ChatGPT [*]	评价 ChatGPT 生成护理计划文本的效果	ChatGPT 生成的文本在可靠性、护理信息质量和整体质量方面表现一般
Musallam 等 ^[36]	美国	未提及	护理教育	ChatGPT [*]	进行护理模拟的新型教学	护生可通过 ChatGPT 集成系统,模拟患者沟通,获取相关的临床信息,但系统存在响应延迟和情感表达不足的问题
Simms 等 ^[37]	美国	未提及	护理教育	ChatGPT [*]	将 ChatGPT 等生成式人工智能融入具体实践教学	可提高护生的学术诚信、信息验证和临床判断力
Lane 等 ^[38]	美国	描述性研究	护理教育	ChatGPT [*] 、Google Bard	介绍生成式人工智能在护理教育中的实施,并为其战略使用提供了指南	生成人工智能既可以作为工具,也可能成为“暴君”,其在提供效率和创新机会的同时,也带来了透明性、伦理使用等挑战
马应卓等 ^[39]	中国	质性研究	护理教育	大语言模型	护生使用大语言模型体验	共提炼 4 个主题,使用体验感多样化、使用场景差异化、使用潜在问题、自我感知需求和期待

注:* 表示未说明 ChatGPT 具体类型。

2.3 大语言模型在护理学领域的研究内容 纳入文献中,8 篇^[7-8,16-18,24,26,34] 文献评估各种大语言模型在护理专业考试或护理基础知识测试的水平,包括执业护士考试;3 篇^[9,23,39] 文献基于访谈内容分析了用户的使用体验,并为整合大语言模型到护理教育和临床护理提供参考依据;7 篇^[10-11,14,20-21,27,35] 文献探索了其

生成护理专业化文本的质量,包括护理教学写作、临床模拟场景构建、制订护理计划等;5 篇^[12-13,22,29,36] 文献讨论了大语言模型在护理教育实际场景的集成与应用,辅助医学知识理解、辅助临床场景学习等,分析了其在教学活动中的具体效果;6 篇^[15,25,28,31-33] 文献探讨了大语言模型在临床护理的集成与应用,包括辅

助护理决策,缓解术前焦虑,协助疾病管理,分析了其在临床护理场景的效果;1 篇^[19]文献探索了护生 ChatGPT 的认知和需求,发现目前中国护生对大语言模型的认知较低;3 篇^[30,37-38]文献对目前大语言模型在护理领域的优势和挑战提出了使用方法、规范、策略。

2.4 大语言模型在护理学领域的应用效果 纳入的文献显示,大语言模型已经在护理教育、临床护理以及患者管理方面展现出了显著的应用效果。在护理教育方面,GPT-4 和 GPT-3.5 表现出了较高的护理知识考试准确率^[7-8],且 ChatGPT 生成的考试题在清晰度、语法、临床相关性和术语复杂度上与护理教育者相当^[12]。此外,ChatGPT 还能帮助护理教育者开发临床模拟场景^[27],通过模拟情境化的患者背景故事,提高学生的参与度、临床判断能力,减少焦虑并提高学习成果^[13],从而达到提升教学效果的目的。部分大语言模型也可作为护生的学习工具^[12],如 ChatGPT 可为护理学生的写作提供详细且人性化的反馈,并提出具体的改进建议^[10]。临床护理和患者管理方面,大语言模型能用于协助护理计划的制订^[21],且搭载 ChatGPT 的大数据随访平台能够提高护理工作效率,协助患者病情管理^[15],并为患者提供专业知识和信息支持^[20-21],大部分患者对 ChatGPT 的相关回答也表现出满意的态度^[32]。然而,大语言模型在护理学领域的应用也存在一些不足之处,如在知识应用、回答临床问题、进行专业判断或决策时,仍不足以取代人工^[14,16-17,24-25,28,34]。

3 讨论

3.1 大语言模型在护理学领域存在研究不均衡现象,且缺乏高质量实证研究 本研究显示,英文文献发文量美国居首位,这可能与目前流行的大语言模型 ChatGPT、Google Bard 等由美国率先发布,且部分国家的护理人员目前无法通过官方网站访问有关。中国学者发文量 13 篇(中文 5 篇,英文 8 篇),居于前列,这表明我国护理学者对大语言模型在护理学相关研究关注度较高。近年来,国内逐渐开展对大语言模型的本土化探索,如文心一言和通义千问等国内易于访问的大语言模型,这类本土化大语言模型在理解中文能力、分析中文语境等方面表现较为出色。本研究纳入文献的研究类型较为丰富,包括描述性研究、实验类研究、回顾性研究、质性研究,但文献类型分布不均衡,多数文献基于描述性研究评估大语言模型在护理学领域的应用潜力,缺乏高质量的实验性研究,以进一步验证大语言模型在护理学领域的应用价值。高质量实证研究的缺乏可能与目前大语言模型在护理学领域处于起步阶段,护理学领域的特殊性、大语言模型的局限性、跨学科难度和技术壁垒等因素相关。

3.2 大语言模型在护理学领域具有较高的应用价值,但局限性与风险性并存 本研究结果显示,大语言模型在护理教育、临床护理、护理科研领域具有较大的应用价值,主要体现在:提高了护理工作和学习的质量和效率;促进了全球护理信息共享;充当了日常学习和办公的工具库(翻译、写代码等)和护理学信息检索库。护理教育方面,通过智能搜索和推荐系统改变了传统护理教学方式,提升了学习和工作的动力和效率,促进了人工智能与护理教育的创新融合,这与李钥等^[40]的研究一致。临床护理方面,大语言模型对护士和患者带来了切实的临床和经济效益,尤其在提高护理效率、优化医疗资源、提升患者满意度方面,具有较大应用前景。护理科研方面,大语言模型推动了护理领域前沿发展和促进学科交叉应用。然而,大语言模型在护理学领域的应用也有其固有的局限,所带来的风险也不容小觑。局限性主要源于模型自身的结构和训练数据的限制,主要包括:①生成护理信息失误(准确性不足、偏见、幻觉、无证据断言);②缺乏护理学科核心价值(无法共情患者、无同理心、不能以患者为中心、护理信息不具有个性化);③网络障碍(访问地域权限、网络连接失败);④技术障碍(部分大语言模型不能分析非文本数据、字数限制、信息滞后);⑤文化不适应(部分大语言模型在特定语言,如英文,表现力才更佳、推荐的护理信息不符合本土国情);⑥成本考虑(部分大语言模型需要付费)。风险性主要与用户的使用方式和目的相关:①使用依赖,丧失原创力;②泄露数据和违反护理伦理;③使用不诚信(作弊、剽窃、抄袭);④未经授权的使用(通过非官方途径访问);⑤过分信任技术判断,忽视专业判断。

3.3 大语言模型在护理学领域应用的对策分析 考虑到大语言模型在护理学领域的潜在应用价值和使用局限性、风险性并存,因此,制定使用规范和指南的意义就显得尤为重要。Blomquist 等^[41]利用全国护理委员会制定的护理委派五权利框架,指导护理学生如何正确使用大语言模型。2023 年 12 月我国互联网信息办公室也颁布了《生成式人工智能服务管理暂行办法》^[42],旨在促进生成式人工智能健康发展和规范应用,但目前仍然缺乏医疗护理领域的相关规范。因此,需要多方联动,整合护理院校、医院、科研机构、人工智能领域的专家制定可实施、具体化的使用指南,规范护生、护理教师、临床护士、护理管理者等护理专业人员使用行为,同时还需要制定相关惩罚机制和检查手段。其次,作为大语言模型的使用主体,护理专业人员应该树立正确的使用态度和观念,并且结合护理专业的伦理准则和实际需求正确评估和应用模型提供的信息和建议。有研究明确指出,仅对人工智能生成的内容保持正确态度是不够的,护理专业人员在不熟悉大语言模型等人工智能产品工作原理、护

理专业知识较弱的情况下,可能无法鉴别其生成的劣质内容,从而造成差错^[43]。因此,需要开设相关培训,帮助了解最新技术和知识,进而推动护理事业的未来发展。

3.4 对未来研究的启示 尽管大语言模型面临应用局限、使用风险和技术障碍等挑战,但其在护理教育、临床护理与护理科研领域展现出的显著效益和巨大发展潜力,预示着这些技术将在护理学的未来发展中扮演重要角色。近年来,我国卫生部门积极倡导护理学与人工智能技术的整合与交叉应用,这一趋势也表明,在大数据时代背景下,大语言模型对护理学的深远影响。因此,如何更好地在实践中融合和释放大语言模型成为我国护理学者亟待深入研究的问题。综合文献和政策分析可知,我国护理学者尚未开展大语言模型在护理学领域的实证研究,并且缺乏相关使用制度与规范的探索。因此,未来研究可聚焦以下内容展开:①进行质性和量性混合研究探索护理人员的使用体验、影响因素等,旨在了解目前我国护理人员使用现状和需求,为制定使用规范,构建干预方案和科技公司完善大语言模型提供参考依据;②研发大语言模型在护理学领域的评估工具,以科学评估大语言模型应用于护理学各领域的使用效果;③加强我国本土化大语言模型实用程序的研究,拓宽其在护理学的应用空间;④开展长期、多中心、大样本的高质量随机对照试验,评估其在护理教育、临床护理、护理科研中的具体效用;⑤构建护理学领域应用大语言模型的培训方案,提升数字素养。

4 小结

本研究依据 Arksey 范围综述的理论框架,总结了国内外大语言模型在护理学领域开展的研究,从发文数量、研究设计、应用场景、研究内容及结果等方面进行了系统的回顾与分析,为相关领域研究提供了基础依据,但本研究未对纳入文献进行质量评价。总体来看,大语言模型在护理学领域的研究还处于起步阶段,虽然我国护理学者发文量居前列,但缺乏实证研究,医疗护理领域也缺乏相关法律法规和规章制度。未来研究应从以上角度出发,探索其在护理学领域的实际应用效果。

参考文献:

- [1] Von Gerich H, Moen H, Block L J, et al. Artificial Intelligence-based technologies in nursing: a scoping literature review of the evidence[J]. *Int J Nurs Stud*, 2022, 127:104153.
- [2] 崔际环. 我国计算机辅助实施责任制护理软件系统研制成功[J]. *护士进修杂志*, 1987, 2(6):21.
- [3] Mitchell M, Krakauer D C. The debate over understanding in AI's large language models[J]. *Proc Natl Acad Sci USA*, 2023, 120(13):e2215907120.
- [4] 中国医院协会信息专业委员会. ChatGPT 在医疗领域的

应用及其挑战概述[EB/OL]. (2023-11-05)[2024-04-06]. <https://www.chima.org.cn/Html/News/Articles/16568.html>.

- [5] OpenAI. Introducing GPT-4o and more tools to ChatGPT free users[EB/OL]. (2024-05-13)[2024-05-20]. <https://openai.com/index/gpt-4o-and-more-tools-to-chatgpt-free/>.
- [6] Arksey H, O'Malley L. Scoping studies: towards a methodological framework[J]. *Int J Soc Res Methodol*, 2005, 8(6):19-32.
- [7] Kaneda Y, Takahashi R, Kaneda U, et al. Assessing the performance of GPT-3.5 and GPT-4 on the 2023 Japanese nursing examination[J]. *Cureus*, 2023, 15(8):e42924.
- [8] Taira K, Itaya T, Hanada A. Performance of the large language model ChatGPT on the national nurse examinations in Japan: evaluation study[J]. *JMIR Nurs*, 2023, 6(1):e47305.
- [9] Chan M M K, Wong I S F, Yau S Y, et al. Critical reflection on using ChatGPT in student learning: benefits or potential Risks? [J]. *Nurse Educ*, 2023, 48(6):E200-E201.
- [10] Parker J L, Becker K, Carroca C. ChatGPT for automated writing evaluation in scholarly writing instruction [J]. *J Nurs Educ*, 2023, 62(12):721-727.
- [11] Cox R L, Hunt K L, Hill RR. Comparative analysis of NCLEX-RN questions: a duel between ChatGPT and human expertise[J]. *J Nurs Educ*, 2023, 62(12):679-687.
- [12] Hsu M H. Mastering medical terminology with ChatGPT and TermBot[J]. *Health Educ J*, 2023; 83(4):352-358.
- [13] Quattrini V, Roesch A, Kretz D. Innovative teaching strategies in DNP education using ChatGPT[J]. *Nurse Educ*, 2024, 49(3):E162-E163.
- [14] Branum C, Schiavenato M. Can ChatGPT accurately answer a PICOT question? assessing AI response to a clinical question[J]. *Nurse Educ*, 2023, 48(5):231-233.
- [15] 夏浩然, 陈小艳, 赵慧明, 等. 搭载 ChatGPT 的人工机械瓣膜置换术后患者自我监测随访大数据平台的建设与实践[J]. *中国实用护理杂志*, 2023, 39(29):2276-2284.
- [16] Huang H. Performance of ChatGPT on registered nurse license exam in Taiwan: a descriptive study[J]. *Healthcare*, 2023, 11(21):2855.
- [17] 刘亚玲, 吕勇, 康晨曦, 等. 消化医护人员与新型聊天机器人在消化内镜癌症筛查相关知识的调查表现[J]. *中华消化内镜杂志*, 2023, 40(11):892-899.
- [18] Wang T, Mu J, Chen J, et al. Comparing ChatGPT and clinical nurses' performances on tracheostomy care: a cross-sectional study[J]. *Int J Nurs Stud Adv*, 2024, 6:100181.
- [19] Luo Y, Weng H, Yang L, et al. College students' employability, cognition, and demands for ChatGPT in the AI era among Chinese nursing students: web-based survey[J]. *JMIR Form Res*, 2023, 7(1):e50413.

- [20] 谭晓文,陈文芳,王娜娜,等. 国内不同大型语言模型对前列腺癌围术期护理与健康教育相关问题的查询响应与效果评价[J]. 中华男科学杂志,2024,30(2):151-156.
- [21] Karacan E. Evaluating the quality of postpartum hemorrhage nursing care plans generated by Artificial Intelligence models[J]. J Nurs Care Qual, 2024, 39(3):206-211.
- [22] Chang C Y, Yang C L, Jen H J, et al. Facilitating nursing and health education by incorporating ChatGPT into learning designs[J]. Educational Technology & Society, 2024,27(1):215-230.
- [23] Alanzi T M. Impact of ChatGPT on teleconsultants in healthcare:perceptions of healthcare experts in Saudi Arabia[J]. J Multidiscip Healthc,2023,16:2309-2321.
- [24] Zong H, Li J, Wu E, et al. Performance of ChatGPT on Chinese national medical licensing examinations: a five-year examination evaluation study for physicians, pharmacists and nurses[J]. BMC Med Educ, 2024, 24(1):143.
- [25] Saban M, Dubovi I. A comparative vignette study:evaluating the potential role of a generative AI model in enhancing clinical decision-making in nursing[J]. J Adv Nurs,2024;Epub ahead of print. DOI:10.1111/jan.16101.
- [26] Su M C, Lin L E, Lin L H, et al. Assessing question characteristic influences on ChatGPT's performance and response-explanation consistency: insights from Taiwan's nursing licensing exam[J]. Int J Nurs Stud, 2024, 153:104717.
- [27] Vaughn J, Ford S H, Scott M, et al. Enhancing healthcare education:leveraging ChatGPT for innovative simulation scenarios[J]. Clin Simul Nurs,2024, 87:101487.
- [28] Zaboli A, Brigo F, Sibilio S, et al. Human intelligence versus ChatGPT: who performs better in correctly classifying patients in triage? [J]. Am J Emerg Med, 2024, 79:44-47.
- [29] Woodley L K. Harnessing the power of Artificial Intelligence to guide nurse educator students[J]. Nurse Educ, 2024,49(5):E292-E293.
- [30] Sblendorio E, Dentamaro V, Lo Cascio A, et al. Integrating human expertise & automated methods for a dynamic and multi-parametric evaluation of large language models' feasibility in clinical decision-making[J]. Int J Med Inform, 2024,188:105501.
- [31] Yahagi M, Hiruta R, Miyauchi C, et al. Comparison of conventional anesthesia nurse education and an Artificial Intelligence Chatbot (ChatGPT) intervention on preoperative anxiety: a randomized controlled trial[J]. J Perianesth Nurs,2024;Epub ahead of print. DOI:10.1016/j.jopan.2023.12.005.
- [32] Dubin J A, Bains S S, Derogatis M J, et al. Appropriateness of frequently asked patient questions following total hip arthroplasty from ChatGPT compared to arthroplasty-trained Nurses[J]. J Arthroplasty,2024,39(9S1):S306-S311.
- [33] Levin C, Kagan T, Rosen S, et al. An evaluation of the capabilities of language models and nurses in providing neonatal clinical decision support[J]. Int J Nurs Stud, 2024,155:104771.
- [34] 穆兰,徐文博,王学通. 类 ChatGPT 大语言模型在护理教育中应用的实证探讨与前景展望[J]. 卫生职业教育: 2024,42:1-4.
- [35] Dağci M, Çam F, Dost A. Reliability and quality of the nursing care planning texts generated by ChatGPT[J]. Nurse Educ,2024;49(3):E109-E114.
- [36] Musallam E, Alhaj Ali A, Alkhafaji M. OpenAI's ChatGPT clinical simulation: an innovative teaching strategy for clinical nursing education [J]. Nurse Educ, 2024; Epub ahead of print. DOI:10.1097/NNE.0000000000001657.
- [37] Simms R C. Work with ChatGPT, Not against: 3 teaching strategies that harness the power of Artificial Intelligence [J]. Nurse Educ,2024,49(3):158-161.
- [38] Lane S H, Haley T, Brackney D E. Tool or tyrant: guiding and guarding generative Artificial Intelligence use in nursing education[J]. Creat Nurs,2024,30(2):125-132.
- [39] 马应卓,王俊,李雪连,等. 护生对大语言模型聊天机器人使用体验的质性研究[J]. 护理学杂志,2024,39(16):73-74.
- [40] 李钥,淮盼盼,杨辉. ChatGPT 在护理教育中的应用状况及优劣分析[J]. 护理学杂志,2023,38(21):117-121.
- [41] Blomquist J, Llewellyn S, Alderden J, et al. Empowering faculty to incorporate large language models in nursing education using a delegation framework[J]. Nurs Educ Perspect,2024;Epub ahead of print. DOI:10.1097/01.NEP.0000000000001246.
- [42] 中央网络安全和信息化委员会办公室. 生成式人工智能服务管理暂行办法[EB/OL]. (2023-07-13)[2024-04-07]. https://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm.
- [43] Adler-Milstein J, Redelmeier D A, Wachter R M. The limits of clinician vigilance as an AI safety bulwark[J]. JAMA,2024,331(14):1173-1174.

(本文编辑 黄辉,吴红艳)