

# 护生对大语言模型聊天机器人使用体验的质性研究

马应卓<sup>1</sup>, 王俊<sup>1</sup>, 李雪莲<sup>1</sup>, 孙文<sup>2</sup>, 吴洁<sup>1</sup>, 刘彤<sup>1</sup>, 赵庆华<sup>1</sup>, 肖明朝<sup>3</sup>

**摘要:**目的 了解护生对大语言模型聊天机器人的使用体验,为规范使用大语言模型聊天机器人提供参考。方法 采用描述性质性研究方法,以目的抽样法选取 15 名来自 4 所医学院校的护生进行半结构式访谈。使用内容分析法对访谈资料进行分析、归纳并提炼主题。结果 共提炼出 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。结论 护理教育和管理者应引领护生树立正确使用大语言模型价值观,制定相关使用准则规范其使用,还须加强培训和护理伦理建设,促进大语言模型在护理教育中的创新融合。

**关键词:**护生; 大语言模型; 聊天机器人; 人工智能; ChatGPT; 数字素养; 护理教育; 质性研究

**中图分类号:**R47;TP242.6 **DOI:**10.3870/j.issn.1001-4152.2024.16.069

## Qualitative study on nursing students' experience in the use of large language model chatbots

Ma Yingzhuo, Wang Jun, Li Xuelian, Sun Wen, Wu Jie, Liu Tong, Zhao Qinghua, Xiao Mingchao. Department of Nursing, The First Affiliated Hospital of Chongqing Medical University, Chongqing 400016, China

**Abstract: Objective** To understand nursing students' experience in the use of large language model chatbots, and to provide a reference for standardized usage of large language model chatbots. **Methods** Descriptive qualitative research approach was adopted. Using purposive sampling strategies, online semi-structured interviews were conducted with 15 nursing students from 4 medical universities. Content analysis method was used to analyze, summarize, and extract themes from the interview data. **Results** A total of 4 themes were extracted: diverse experiences of usage; differentiation of usage scenarios; potential issues of usage; self-perceived needs and expectations. **Conclusion** Nursing educators and managers should lead nursing students to establish correct values towards the usage of large language models, develop relevant guidelines to standardize their usage, and must strengthen training and nursing ethics construction, in an effort to promote the innovative integration of large language models in nursing education.

**Keywords:** nursing students; large language model; chatbot; artificial intelligence; ChatGPT; digital literacy; nursing education; qualitative research

人工智能的快速发展促进了基于大语言模型的聊天机器人出现和不断更新迭代。此类聊天机器人是指包含数十至万亿参数的 Transformer 语言模型,其使用深度学习等技术,并接受大量文本数据训练,能够完成自然语言处理和人机交互等复杂任务<sup>[1-2]</sup>。ChatGPT 是目前最流行的大语言模型聊天机器人之一,2022 年一经发布便因其理解上下文、生成连贯的文本,拒绝不适当的请求等功能受到社会广泛关注,成为历史上增长速度最快的 App<sup>[3-5]</sup>。与此同时,国内外各大科技公司也陆续研发大语言模型,如我国研发的文心一言、讯飞星火认知大模型等<sup>[6]</sup>。大语言模型因其出色的人机对话和互动能力,在计算机、新闻及医疗护理等方面展现出广阔的应用和研究前景。相关研究表明,大语言模型聊天机器人具有为护生提供个性化的学习内容、模拟临床护理场景、减轻学习压力等优势,备受护生关注和喜爱<sup>[7-8]</sup>。但其在使用

过程中仍存在诸多问题,主要包括大语言模型在护理场景中的应用缺陷(专业性不足、可借鉴性不高、误导性强)和护生使用风险(护理伦理、患者安全、使用诚信)<sup>[9-12]</sup>。护生作为护理教育的主要参与者,其处理可用资源和使用工具的方式会影响学业成绩和未来执业能力,进而对护理教育模式造成影响。因此,了解护生使用体验及场景,分析大语言模型聊天机器人对护理教育的影响十分必要。目前国内关于大语言模型在护理教育领域中的研究多为综述和现状调查,未深度挖掘用户使用体验。张海等<sup>[13]</sup>通过质性访谈讨论 ChatGPT 用户使用意愿的影响因素,虽然该研究纳入医学相关背景人群,但未深入分析包括护生在内的医学生这一特定群体的主观感受和现存问题。本研究通过质性访谈,剖析和解读护生使用大语言模型的体验及存在的问题,为护理教育及管理者的构建培训方案和制定应对策略提供参考。

## 1 对象与方法

**1.1 对象** 2023 年 9—11 月,采用目的抽样法选择我国 4 所医学院校(浙江大学、重庆医科大学、贵州中医药大学、西南医科大学)的护生作为研究对象,进行线上半结构式访谈。纳入标准:自愿参与本研究;使用过或正在使用基于大语言模型的聊天机器人。排除标准:因各种原因办理休学、退学护生。以信息饱和(收集的资料内容重复且资料分析无新的主题出

作者单位:重庆医科大学附属第一医院 1. 护理部 3. 泌尿外科(重庆, 400016);2. 四川大学华西医院

马应卓:男,硕士在读,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);中华医学会医学教育分会医学教育研究重点课题(2020A-06044);重庆市高等教育教学改革研究项目重点项目(222069)

收稿:2024-04-07;修回:2024-05-30

现)为原则<sup>[14]</sup>,共访谈 15 名护生,编号 N1~N15;其中女 11 名,男 4 名;年龄 19~32(24.27±4.30)岁。博士研究生 4 名,硕士研究生 6 名,本科生 5 名。

## 1.2 方法

**1.2.1 确定访谈提纲** 根据研究目的进行文献回顾和分析,结合课题组讨论意见形成访谈提纲初稿,咨询 2 名从事人工智能和医疗交叉领域的专家(1 名研究人工智能和医学交叉领域的临床医学博士,1 名研究医疗智能服务系统的护理学博士),根据咨询意见修改提纲,并在重庆医科大学选取 3 名护生进行预访谈,根据预访谈结果及受访者意见修订形成正式提纲:①请您谈谈对大语言模型聊天机器人的认识与理解;②请您回忆并描述最近 1 个月内使用大语言模型聊天机器人过程中印象深刻的事件;③您在使用过程中存在哪些积极和消极体验?④您觉得使用大语言模型聊天机器人会造成哪些影响?⑤您的主要用途是什么?⑥还有什么使用经验或建议想要分享吗?

**1.2.2 资料收集** 采用描述性质性研究方法,对 15 名护生采用线上方式(腾讯会议)进行半结构式深度访谈。向受访者解释本次访谈的主要内容及时长,告知其需要全程录音,承诺访谈相关资料仅用于学术研究,不会暴露个人隐私。征得访谈对象同意后,约定访谈时间并发送会议链接,访谈时间控制在 20~30 min。访谈过程中研究者保持客观态度,避免暗示性提问并及时记录。本研究已获得重庆医科大学附属第一医院伦理审查委员会批准(科研伦理 2023-334 号)。

**1.2.3 资料整理与分析** 访谈结束后 24 h 内将会议录制内容转录为文字,并记录受访者语气停顿、声调变化等。使用内容分析法对访谈资料进行分析:团队 2 名研究者反复阅读资料,并沉浸在资料中,以获得对访谈资料的全面理解;标注出与使用大语言模型相关的重要经历和感受的句子,开始开放性编码;将相似和相关的编码归类形成同类主题;定义主题和编码,并从资料中找出相应的文本<sup>[15]</sup>。共析出 4 个主题。

## 2 结果

### 2.1 使用体验感多样化

**2.1.1 使用获益感** 部分护生表示,大语言模型作为一种人工智能辅助工具对自己的学习和工作是有益的,可以提高学习和工作的效率和质量。N15:“我之前用文心一言帮我梳理零散的护理知识点并归类,随后我完善了一下,我觉得这提高了我的学习效率。”N5:“我论文完成后用它进行润色和查找错别字,这样我有更多的时间去构思我的文章,整体上提升了我的写作质量和效率。”N7:“遇到复杂的护理案例分析,ChatGPT 可以帮助我分析这些材料而且还有解析。”

**2.1.2 使用趣味感** 大语言模型聊天机器人作为人工智能领域内一项新兴前沿技术,通过复杂的自然语言处理和交互式人机对话功能,丰富了护生学习过程中的趣味性和互动性。N2:“当第一次使用 ChatGPT 时,发现它比我想的更加智能化,就好像能听懂并且理解我表达的内容,我觉得它的回答挺有趣的。”此外,包括 ChatGPT4 和文心一言在内的大语言模型具有图片生成功能,能够根据护生的请求生成相关的医学图像,为护生提供全新的视觉学习方式,该方式通过增强学习趣味性而激发护生学习的积极性。N6:“一次课堂作业,它根据我的提示生成了一张甲状腺的解剖图片,而且它可以根据我的提示更改,相比之下它比一些资料书更加智能有趣!”

**2.1.3 使用倦怠感** 部分护生表示大语言模型不能达到自己的预期目标,存在使用倦怠感,这种消极体验归因于大语言模型在运用中所暴露出的缺陷:护理专业知识局限性、生成内容普适性、信息滞后性。N6:“刚开始感觉还挺智能的,后来发现对于一些专业性问题表现不好,还需要去核查。”N8:“因为 ChatGPT 本身生成的信息不具有时效性,毕竟医学知识需要不断更新,所以用得也少了。”N4:“我有时候感觉它在糊弄我,就是回答的护理措施很大众化,不具有针对性。”

**2.2 使用场景差异化** 护生在大语言模型的应用场景和使用偏好存在明显差异性,这种差异性与他们的大语言模型聊天机器人的多元认知、现阶段需求、个人兴趣、数字素养有关。N1:“我对它还是比较感兴趣,觉得功能很强大,我一般用于辅助科研比较多,而且我认为它对某些疾病预测肯定也是有帮助的,相关文献也有过报道。”N4:“我认为它对我帮助挺大的,我一般会问它一些临床护理知识。”N10:“我不太会在科研中使用它,因为这非常不科学,况且现在我才大四,没有写论文这方面的需求。”N8:“我现阶段就主要将它用来翻译英文,其实有没有都差不多,其他翻译软件也可以代替它,其次,我在计算机和人工智能方面是个小白,不想搞那么复杂。”

### 2.3 使用潜在问题

**2.3.1 过度依赖** 大语言模型以迅速提供相对准确的信息和答案的方式,加剧护生的惰性思维,降低了学习热情和独立解决问题的能力。N7:“我觉得我有点过度依赖它了,有些时候就觉得它很方便,自己也不用太思考。”N4:“我觉得我简直离不开它了,它给我学习带来了非常大的帮助。”

**2.3.2 不规范应用** 部分护生表示,他们曾经使用过大语言模型代替自己完成老师布置的课后作业,甚至有护生在护理基础知识测试中和论文写作中使用大语言模型。N15:“平时一些心得体会、思想感悟之类的作业我就会立刻想到让它帮助我完成。”N11:“一次线上考试,因为我没太大把握,就用了文心一言

去搜题。”N9：“我认识的一个护理研究生，在一次论文作业中，就使用了类似 ChatGPT 的机器人替写。”

**2.3.3 数据安全意识不足** 护生在使用过程中个人数据保护意识弱。N12：“我并不觉得会造成数据泄露，自己的信息也没什么可值得泄露的。”此外，部分护生在好奇心的驱使下会探索大语言模型的临床护理能力，这个过程存在泄露患者隐私的风险。N8：“因为比较好奇，我把一些患者的诊疗信息输入到 ChatGPT 中，想看看它的疾病分析能力，结果表现一般。”

## 2.4 自我感知需求和期待

**2.4.1 开设相关课程** 部分护生对大语言模型相关知识掌握程度和应用能力处于较低水平，这限制了他们发挥大语言模型的潜力，他们希望学校能够提供相关培训、开设相关课程。N14：“大语言模型聊天机器人回答质量的优劣在一定程度上取决于问答者的提问方式和聊天技巧，所以我挺希望学校有这方面的讲座或者课程教我们如何使用。”N13：“我其实不太了解它的使用原理，还有它的许多潜在功能也不知道，希望学校有这方面的培训。”

**2.4.2 开发和完善医学相关的大语言模型** 针对目前大语言模型在护理领域应用中的缺点，护生认为需要开发针对医疗护理方面的大语言模型，这类模型需要使用科学性强和质量性高的临床证据去训练，确保生成内容可靠并且符合临床应用标准。N3：“我觉得现有的大语言模型聊天机器人在医学和护理方面应用不太可行，其智能程度还是远远不够，应该在现有基础上运用医学和护理专业知识或高质量文献去训练它，形成一个医学大语言模型聊天机器人。”N7：“像现在 ChatGPT4 和文心一言可以根据提示生成一些解剖结构和临床场景的图片，但是准确度不够，我挺希望这方面可以加强一下。”

## 3 讨论

**3.1 树立正确使用观念及态度** 护生对大语言模型聊天机器人的体验感不尽相同，使用过程中正面和负面感受共存。多数护生的使用体验感较好，倾向使用。但是他们也认识到现有的大语言模型聊天机器人并不是针对护理学科而设计，缺乏系统的护理专业知识背景，并不能在一些专业化场景中呈现良好表现，还存在信息泄露、无证据断言等风险。因此，护生需要树立正确的使用价值观念，批判性地对待大语言模型生成的内容，并查阅资料核实。同时，护生需要明确技术与学习之间的关系，合理使用大语言模型等人工智能产品，避免过度依赖和滥用，这与李钥等<sup>[16]</sup>研究结果一致。护理教育者应以发展的眼光看待新技术，紧跟信息化潮流，把握护理教育学科发展趋势。将 ChatGPT 等大语言模型有效整合到护理教育中，促进护理学科的创新发展，护生、

护理教育者和管理者树立正确使用观念及态度是关键一步。

**3.2 开展应用培训，提升护生数字素养** 本访谈结果显示，护生对大语言模型存在认知不足、使用熟练度较低，甚至存在技术恐惧等问题，这与部分护生数字素养不足有关。马亚男等<sup>[17]</sup>通过问卷调查以自我评估方式分析了医学生对大语言模型的了解程度，其中各年级医学生对其了解程度均处于较低水平，本科一年级医学生自我评估“相当了解”部分仅占 1.52%。教育数字化是我国开辟教育发展新赛道和塑造教育发展新优势的重要突破口<sup>[18]</sup>。因此，需要构建高效科学的数字化培训体系和方案破除护生“信息茧房”，提升数字素养，强化护生大语言模型相关知识和应用能力。院校可通过开设大语言模型相关的课程、知识讲座及工作坊等加强培训，护生也应该注重多学科的交流与沟通，主动与计算机科学、人工智能等领域的学生进行科研和日常学习合作，以提升数字素养。

**3.3 制定相关标准，规范护生使用行为** 本研究结果显示，部分护生在使用大语言模型聊天机器人过程中存在过度依赖、数据保护意识薄弱等问题。美国一项调研发现，89% 的学生用 ChatGPT 写作业，并且 48% 的学生承认在考试中使用 ChatGPT<sup>[19]</sup>，该研究引发人工智能技术在教育领域的激烈讨论，研究者认为制定相关管理标准、规范使用者行为迫在眉睫。因此，护理教育者和管理者需与人工智能、机器学习、自然语言处理等相关专业人员共同制定使用准则，规范和引导护生正确使用大语言模型。

**3.4 防范技术风险，加强护理伦理建设** 由于人工智能技术参与医疗护理中的伦理界限不够清晰，护生为减轻学习和工作压力直接使用大语言模型进行日常护理活动，可能会损害患者利益或安全，甚至造成护理不良事件，引发护理伦理问题。医院管理者需加强患者信息管理，严格明确大语言模型在诊疗护理过程中的使用界限，提高数据保护意识，保护患者隐私和医疗安全。

## 4 结论

本研究结果显示，护生对大语言模型聊天机器人使用意愿较高，但目前将其应用于护理教育存在一定风险。护理教育者和管理者应引领护生树立正确使用大语言模型价值观，制定相关使用准则规范其使用，还须加强培训和护理伦理建设，促进大语言模型在护理教育中的创新融合。本研究的对象局限于护生，后续还需从其他专业的医学生以及护理教师、临床医务工作者的视角出发，探索其使用体验感，深入了解具有医学背景的相关人员对大语言模型的使用态度和行为，为制定大语言模型聊天机器人在医疗场景中的使用规范、组织相关培训和课程提供更全面的依据。

参考文献:

[1] Thirunavukarasu A J, Ting D, Elangovan K, et al. Large language models in medicine[J]. *Nat Med*, 2023, 29(8): 1930-1940.

[2] Mitchell M, Krakauer D C. The debate over understanding in AI's large language models[J]. *Proc Natl Acad Sci U S A*, 2023, 120(13): e2079060176.

[3] Teubner T, Flath C M, Weinhardt C, et al. Welcome to the era of ChatGPT[J]. *Bus Inf Syst Eng*, 2023, 65(2): 95-101.

[4] Thorp H H. ChatGPT is fun, but not an author[J]. *Science*, 2023, 379(6630): 313.

[5] Gordon C. ChatGPT is the fastest growing App in the history of Web Applications[EB/OL]. [2023-02-28]. <https://www.forbes.com/sites/cindygordon/2023/02/02/chatgpt-is-the-fastest-growing-ap-in-the-history-of-web-applications/>.

[6] 令小雄,王鼎民,袁健. ChatGPT 爆火后关于科技伦理及学术伦理的冷思考[J]. *新疆师范大学学报(哲学社会科学版)*, 2023, 44(4): 123-136.

[7] Berse S, Akca K, Dirgar E, et al. The role and potential contributions of the artificial intelligence language model ChatGPT[J]. *Ann Biomed Eng*, 2024, 52(2): 130-133.

[8] Liu J, Liu F, Fang J, et al. The application of Chat Generative Pre-trained Transformer in nursing education[J]. *Nurs Outlook*, 2023, 71(6): 102064.

[9] 郭彩霞,郭彩旭,史晓宁,等. ChatGPT 赋能护理实践:前景、风险及对策[J]. *协和医学杂志*, 2023, 14(6): 1170-1174.

[10] Seney V, Desroches M L, Schuler M S. Using ChatGPT to teach enhanced clinical judgment in nursing education[J]. *Nurse Educ*, 2023, 48(3): 124.

[11] 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.

[12] Dave T, Athaluri S A, Singh S. ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations[J]. *Front Artif Intell*, 2023, 6: 1169595.

[13] 张海,刘畅,王东波,等. ChatGPT 用户使用意愿影响因素研究[J]. *情报理论与实践*, 2023, 46(4): 15-22.

[14] 李峥,刘宇. *护理学研究方法*[M]. 北京:人民卫生出版社, 2018: 245.

[15] 朱丹玲,周云仙. 护理文献中质性内容分析法的应用误区探讨[J]. *护理学杂志*, 2018, 33(21): 97-99.

[16] 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.

[17] 马亚男,杨瑞丰,何小东. 类 ChatGPT 大规模语言模型对医学生的影响[J]. *医学教育研究与实践*, 2023, 31(5): 526-532.

[18] 中华人民共和国教育部. 加快推进教育数字化建设教育强国 [EB/OL]. (2022-12-09) [2024-01-01]. [http://www.moe.gov.cn/jyb\\_xwfb/s5148/202212/t20221209\\_1028299.html](http://www.moe.gov.cn/jyb_xwfb/s5148/202212/t20221209_1028299.html).

[19] Study.com. Productive teaching tool or innovative cheating? [EB/OL]. (2023-01-30) [2024-01-01]. <https://study.com/resources/perceptions-of-chatgpt-in-schools>.

(本文编辑 丁迎春)

(上接第 59 页)

[15] 屈歌,雷子辉,冯晶,等. 中国急诊科护士职业倦怠现状及其空间分布研究[J]. *中国卫生政策研究*, 2022, 15(10): 71-76.

[16] Karakoc A, Yilmaz M, Alcalar N, et al. Burnout syndrome among hemodialysis and peritoneal dialysis nurses[J]. *Iran J Kidney Dis*, 2016, 10(6): 395-404.

[17] 曾凡殊,刘霜莹,贺英. 急诊科护士职业认同感、职业倦怠感现状及综合管理对策[J]. *职业与健康*, 2022, 38(8): 1126-1129, 1134.

[18] de la Fuente-Solana E I, Suleiman-Martos N, Pradas-Hernandez L, et al. Prevalence, related factors, and levels of burnout syndrome among nurses working in gynecology and obstetrics services: a systematic review and meta-analysis[J]. *Int J Environ Res Public Health*, 2019, 16(14): 2589.

[19] 谷娅. 三甲医院重症监护室护士主观幸福感与心理资本、工作压力相关性[J]. *中国健康心理学杂志*, 2019, 27(2): 264-267.

[20] 万娅姣,王英,夏添. 护士心理资本和职业认同的关系研究[J]. *护理学杂志*, 2013, 28(3): 56-58.

[21] 高怡明,赵香玉,王清,等. 手术室护士心理资本对职业

倦怠的影响:工作沉浸的中介作用及性别调节[J]. *山东大学学报(医学版)*, 2023, 61(1): 74-79.

[22] 孙明华,赵静静,郭传敏. 影像科护士共情能力与职业倦怠和主观幸福感的关系[J]. *工业卫生与职业病*, 2019, 45(5): 373-376.

[23] Thirioux B, Birault F, Jaafari N. Empathy is a protective factor of burnout in physicians: new neuro-phenomenological hypotheses regarding empathy and sympathy in care relationship[J]. *Front Psychol*, 2016, 7: 763.

[24] Williams B, Lau R, Thornton E, et al. The relationship between empathy and burnout-lessons for paramedics: a scoping review[J]. *Psychol Res Behav Manag*, 2017, 10: 329-337.

[25] Franco P, Tesio V, Bertholet J, et al. The role of alexithymia and empathy on radiation therapists' professional quality of life[J]. *Tech Innov Patient Support Radiat Oncol*, 2020, 15: 29-36.

[26] Salvarani V, Rampoldi G, Ardenghi S, et al. Protecting emergency room nurses from burnout: the role of dispositional mindfulness, emotion regulation and empathy[J]. *J Nurs Manag*, 2019, 27(4): 765-774.

(本文编辑 赵梅珍)