

基于大语言模型的个体预后或诊断多变量模型 透明报告指南:TRIPOD-LLM的解读

潘新伟¹,陈红爽¹,梁诗聪²,王卿驰¹,王亚琦¹,宁时安¹,徐雪蕾¹,史纪元³,霍晓鹏¹

(1.中国医学科学院北京协和医院 护理部,北京 100730;

2.中国医学科学院北京协和医学院 护理学院,北京 100144;

3.北京中医药大学 护理学院,北京 100029)

【摘要】 随着大语言模型(large language models,LLMs)在医学诊断、科研与教育等领域的广泛应用,其卓越的生成与推理能力已显著展现优势。然而,医学领域对伦理、隐私保护及模型准确性的严格标准,也使LLMs的实际应用面临严峻挑战。尽管个体预后或诊断的多变量模型透明报告(transparent reporting of a multivariable prediction model for individual prognosis or diagnosis,TRIPOD)+人工智能(artificial intelligence,AI)为预后或诊断预测模型提供了报告规范,但其在生成式人工智能研究中的适用性仍显不足。本文解读了在TRIPOD+AI基础上扩展形成的TRIPOD-LLM报告指南,系统梳理了其在模型构建、验证、任务适应性及人类监督等方面的报告要素,为提升医学领域LLMs研究的透明度、规范性与可复现性提供了参考。

【关键词】 大语言模型;解读;报告规范;TRIPOD-LLM;预测模型

doi:10.3969/j.issn.2097-1826.2025.11.019

【中图分类号】 R47 【文献标识码】 A 【文章编号】 2097-1826(2025)11-0078-05

Interpretation of TRIPOD-LLM: A Guideline for Transparent Reporting of Multivariable Models for Individual Prognosis or Diagnosis Based on Large Language Models

PAN Xinwei¹, CHEN Hongshuang¹, LIANG Shicong², WANG Qingchi¹, WANG Yaqi¹, NING Shian¹, XU Xuelei¹, SHI Jiyuan³, HUO Xiaopeng¹ (1. Nursing Department, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing 100730, China; 2. School of Nursing, Peking Union Medical College, Chinese Academy of Medical Sciences Beijing 100144, China; 3. School of Nursing, Beijing University of Traditional Chinese Medicine, Beijing 100029, China)

Corresponding author: HUO Xiaopeng, Tel: 010-69155819

【Abstract】 With the widespread application of large language models (LLMs) in medical diagnosis, scientific research, education, and other fields, their outstanding generative and reasoning capabilities have demonstrated significant advantages. However, the strict standards for ethics, privacy protection, and model accuracy in the medical field have also posed severe challenges to the practical application of LLMs. Although the *Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis* (TRIPOD) + Artificial Intelligence (AI) provides reporting norms for prognosis or diagnosis prediction models, its applicability in generative artificial intelligence research is still insufficient. This article interprets the *TRIPOD-LLM* reporting guideline, which is extended based on TRIPOD+AI. It systematically sorts out its reporting elements in terms of model construction, validation, task adaptability, and human supervision, providing an important reference for improving the transparency, standardization, and reproducibility of LLM research in the medical field.

【Key words】 large language model; interpretation; reporting guideline; TRIPOD-LLM; prediction model

[Mil Nurs, 2025, 42(11): 78-82]

大语言模型(large language models,LLMs)在

【收稿日期】 2025-07-16 **【修回日期】** 2025-10-21
【基金项目】 中央高水平医院临床科研业务(2022-PUMCH-B-130)
【作者简介】 潘新伟,本科,主管护师,电话:010-69155819
【通信作者】 霍晓鹏,电话:010-69155819

医学研究中的应用正蓬勃兴起,尤其在护理领域展现较大潜力。它们能够辅助构建精准预测模型,为临床决策提供科学的数据支持^[1-2],从而为护理研究的创新发展提供坚实的技术支撑^[3-4]。目前,LLMs常应用于支持疾病诊断^[5]、医患辅助决策^[6]和慢病

管理^[7],但由于其在医疗领域应用时面临的幻觉问题^[8]、数据隐私保护^[9]以及结果可解释性等问题^[10],其大范围应用仍需深入研究。个体预后或诊断的多变量模型透明报告(transparent reporting of a multivariable prediction model for individual prognosis or diagnosis, TRIPOD)为开发或评估预测模型性能的研究提供了规范化的报告建议^[11]。为了规范 LLMs 在医学中的使用,Gallifant 等^[12]通过多学科专家的德尔菲法和共识会议,制订了 TRIPOD-LLM 报告指南。TRIPOD-LLM 作为 TRIPOD+人工智能的延伸,针对其在医学研究与临床实践中的应用特点,提供了更为系统、透明且严谨的报告规范^[13]。本文旨在解读 TRIPOD-LLM 条目,为推动 LLMs 在护理领域的应用提供方法学支持。

1 TRIPOD-LLM 报告指南内容

现有诊断和预后模型的报告规范尚未涵盖 LLMs 和聊天机器人相关研究中的超参数选择、预测可变性、语言输出评估等关键环节,而这些因素直接影响模型的可靠性。因此,TRIPOD-LLM 应运而生。TRIPOD-LLM 包含 7 个部分,共 19 个条目,其中 LLMs 代表大语言模型和聊天机器人。为适应医疗研究设计与 LLMs 任务的多样性,该规范引入了模块化研究任务结构,对不同研究和任务类型进行系统规范。部分条目仅适用于特定研究或应用,但并不互斥,研究者需结合研究背景确定 LLMs 类型。该规范将随 LLMs 的发展不断更新,可在 <https://tripod-llm.vercel.app/> 获取最新版本。

2 TRIPOD-LLM 报告规范要点解读

本文以 2025 年 1 月发表的“Detection of gastrointestinal bleeding with large language models to aid quality improvement and appropriate reimbursement”^[14] 为例,对条目进行解读。

2.1 标题(条目 1) 标题应清晰表明研究的性质,即是开发、调整还是评估 LLMs,同时明确指出目标人群和结局指标,以便迅速识别研究的适用范围和目的。示例分析:例文题目符合要求,明确指出利用 LLMs 作为工具,辅助识别胃肠道出血(gastrointestinal bleeding, GIB)并助力医疗质量提升与合理报销。

2.2 摘要(条目 2) 摘要部分需在研究目的、方法、结果等方面简洁全面地呈现 LLMs 相关内容。示例分析:例文为结构式摘要,详细报告了 LLMs 在训练集上的模型构建与性能及其任务应用,但未说明所用 LLMs 的名称和版本。

2.3 前言(条目 3~4)

2.3.1 背景(条目 3a~3b) 研究需详细说明 LLMs

应用背景或案例。若已有 LLMs,应解释再次开发的原因,若是评估研究需给出评估理由并引用相关模型结果,新开发 LLMs 则应描述潜在应用场景。还应报告研究对象、预期用途、使用节点及目标用户,并明确说明 LLMs 是取代还是结合现有的金标准。示例分析:例文指出 GIB 诊断代码的敏感性低,而既往相关模型缺乏特异性,因而开发 LLMs 以自动监测 GIB 并辅助报销。研究对象为 GIB 患者,预期用途为临床监测与编码。在方法部分提及与金标准的融合。

2.3.2 目标 描述研究的所有目标,包括 LLMs 的初始开发、微调、评估性能或验证等。示例分析:例文提出开发自然语言处理(natural language processing, NLP)流程,以 LLM 识别 GIB 体征,并展示两个应用场景:(1)开发并验证自动检测 GIB 复发出血的算法;(2)开发并验证用于准确按疾病诊断相关分组编码(medicare severity diagnosis related groups, MS-DRGs)的算法。

2.4 方法(条目 5~15)

2.4.1 数据来源(条目 5a~5e) 模型性能高度依赖数据^[10],LLMs 研究需详细报告数据来源与类型。随着 LLMs 由单模态向多模态发展,训练数据已从文本扩展至音频、图像和视频^[15]。研究需分别描述训练、调整、评估数据集的数据来源,使用这些数据的理由(如网络语料库、临床研究/试验数据、电子健康记录数据)。若使用现存数据,应说明理由、适用性和代表性,涉及合成数据需提供方法与细节。研究还应说明数据性质(定性/定量)、时间范围、预处理流程(清洗、质量核查、缺失值处理等),并在已有 LLMs 研究中引用或声明开发者提供的数据信息。示例分析:例文报告了 2014 年 7 月至 2023 年 6 月的回顾性临床数据,按医院划分训练与验证集,并人工审查护理记录。但未说明缺失和不平衡数据的处理方式。

2.4.2 数据分析方法(条目 6a~6e) 数据分析部分应详细报告名称、版本、最后训练日期,以及生成文本、初始输出与后处理方式等信息。研究者需提供各类分类的详细信息和依据,若适用,还应说明如何确定概率和识别阈值。研究者还应说明 LLMs 的开发与微调过程,包括架构、训练、对齐策略、超参数设置等,并明确输出类型;同时,提供后处理方法及代码,以确保分析方法的透明性和可重复性。示例分析:例文开发了 4 个 NLP 流程,包括 RegEx、MedSpaCy、LLaMA2 微调版及其混合方法,用于修正常见 LLMs 错误,如去除否定短语,并将处理结果转换为二进制分类输出。

2.4.3 LLM 的输出(条目 7a~7e) LLM 输出需涵盖质量指标、结果与临床相关性、评估者信息及性能对比,使读者能够捕捉生成输出的质量指标,如与金标准的一致性、相关性、准确性以及错误的类型和情况。质量评价应多维度的量化方法,如稳定性、与临床场景匹配度、混淆矩阵或校准曲线。若 LLM 应用于临床决策,研究需验证预测与临床结局的相关性,并通过人工评估防止幻觉内容。需明确后处理方法、术语标准化及模型的推理日期。此外,还应报告评估者的程序、背景及一致性,以保证专业性。性能对比则需详细说明与其他模型的比较标准和过程。示例分析:例文展示了模型性能对比并进行外部验证,说明了临床实用性,且通过手动审查优化提示。但未报告评估者信息,缺乏模型表现与人类生成结果的比较说明。

2.4.4 标注(条目 8a~8c) 若研究涉及标注,应说明标注方式、并提供标注指南和参考材料;同时,需报告标注者数量、数据分配比例及一致性指标(如 Cohen's kappa)。此外,对 LLMs 生成内容的标注能够反映标注者的价值偏好^[16],为确保标注过程的透明性,还应提供标注者的专业背景和经验等信息。示例分析:例文中描述通过人工审核确定每份记录中是否存在护理报告的黑便、便血、呕血或其组合的金标准标签。但未说明人工标注的研究者数量、专业背景及数据比例。

2.4.5 提示(条目 9a~9b) 若研究涉及使用提示,需详细描述初始提示语集的创建方法、完善和筛选标准、最终选择与测试方式等内容;同时,说明所用数据集或信息来源、性能评估方式,并说明开发与评估数据集是否存在重叠。示例分析:例文利用提示语引导 LLMs 识别护理记录中的 GIB 症状,明确症状定义及排除规则,并标注相关短语。数据分为训练集、内部验证集和外部验证集,采用正则表达式(RegEx)与 LLMs 结合的混合流程,将二进制输出整合为集成算法,以平衡准确性与敏感性。该研究较完整展示了提示的设计与标准。

2.4.6 预处理(条目 10) 描述在数据汇总前进行数据去识别化等预处理步骤,例如,是否进行了重新格式化或额外处理的文本或章节。示例分析:例文主要任务是从数据集中自动识别标准化文本,不涉及文本的简化等。

2.4.7 指令调整/对齐(条目 11) 若采用了指令调整或对齐策略,需描述在该过程中提供给 LLMs 的具体指令或偏好数据集;同时,需评估人员的人口统计学信息、专业知识情况,以及其用于评估 LLMs 性能并反馈的界面或工具。示例分析:例文未包含自

行进行指令微调或对齐操作,因此该条目不适用。

2.4.8 计算资源(条目 12) 报告研究所需的计算资源或其代理指标,包括机器数量、计算时间、总成本、模型推理时间、每秒浮点运算次数等数据。示例分析:例文未进行相关数据报告,因此难以评估研究的计算成本。

2.4.9 伦理(条目 13) LLMs 的应用涉及患者的数据隐私等问题,需详细报告批准研究的机构研究委员会或伦理委员会的名称,并说明参与者知情同意书或伦理委员会豁免知情同意书的详细情况。示例分析:例文明确报告了研究得到了耶鲁大学医学院机构伦理审查委员会的批准(IRB:1408014519),并说明了数据来源为回顾性数据。

2.4.10 开放科学(条目 14a~14f) 在研究相关信息披露方面,需全面且细致。(1)资金来源及资助者在设计、数据、分析与发表中的角色;(2)全体作者利益冲突与财务披露;(3)研究方案获取途径;(4)若为临床试验,提供注册机构与注册号;(5)数据可用性;(6)可复现结果的代码信息与获取方式或地点。示例分析:例文在文末注明基础设施支持来源,并列作者贡献;补充材料提供部分代码与提示审查改进流程,利于复现。但未见利益冲突声明、资源可及性与研究注册信息,削弱了研究的透明性与可解释性。

2.4.11 公众参与(条目 15) 详细说明患者和公众在研究设计、实施、报告、解释或传播各阶段的参与详情,或明确指出其未参与的情况,并解释是否以及如何将研究结果与患者或公众进行分享。示例分析:例文未提及有关公众参与的情况和说明。

2.5 结果(条目 16a~18)

2.5.1 研究对象(16a~16d) 在使用患者/电子健康记录数据时,描述文本/电子健康记录/患者数据在研究中的流动情况,明确每个分析中的参与者数量和结局事件。具体包括:具有和未具有结局/标签的文档/问题/参与者的数量,以及适用时的随访时间。若研究涉及临床结局相关 LLMs 评估,还需报告开发与评价数据集的人口学特征、预测因素及临床分布等的可比性。示例分析:例文基于电子病历进行识别与提取,但未用流程图展示各阶段的数据处理(16a)。在结果中,提供了样本基线特征及训练/验证/外部验证集的分层比较(16b、16c)。然而,未说明不同阶段样本量的变化或调整(16d)。

2.5.2 性能(条目 17) 描述模型性能有助于评估其在处理数据和生成预测值时所消耗的计算资源,并判断其性能是否满足研究需求^[17]。应根据预先设定的指标和/或人工评估报告来衡量 LLMs 的性能。报告中需涵盖整体及关键亚组的社会人口统计、

诊断、置信区间等性能信息。此外,还应提供 LLMs 语言化估计值等生成结果的不确定性评估,以便读者能够评估模型预测的不确定性。示例分析:例文详细的报告了四种模型的性能对比结果,包含准确率、敏感度、受试者工作曲线下面积等详细数据。

2.5.3 LLMs 的更新(条目 18) LLMs 的更新可能会影响模型性能退化或训练数据出现偏差等情况,若研究中存在 LLMs 的更新,应详细描述对 LLMs 所做的任何修改或更新及其背后的原因,并报告更新后的性能指标。示例分析:例文未报告使用的 LLMs 是否存在更新。

2.6 讨论

2.6.1 阐释(条目 19a) LLMs 在临床实践中用途多样^[18-19],需总结主要研究结果及其总体和对特定或预期的医疗保健使用场景的影响。讨论研究出现的预测中的偏差等任何公平性或稳健性问题。示例分析:例文验证了使用基于 LLMs 的 NLP 管道识别 EHR 文档中 GIB 的表现,可以使 MS-DRG 编码算法能够更准确地识别 GIB 患者适当的 MS-DRG,进行医疗报销资源的合理分配。

2.6.2 局限性(条目 19b) 识别并解释该研究设计的局限性、结果的稳健性,讨论本研究的局限性及其对偏倚、统计不确定性和普遍性的影响,以及这些研究结果在推广应用方面的意义。描述在指定任务和领域背景下使用数据时已知的任何挑战,包括数据表示、缺失数据、数据协调和偏倚方面的问题。示例分析:例文的局限性主要为模型的推广性,以及临床数据中有关重症 GIB 合并症相关描述较多,一定程度上影响文本内容的提取,仍需在一般人群中验证模型的实用性。

2.6.3 LLMs 在应用场景中的可用性(条目 19c~19g) 需全面说明 LLMs 在应用场景中的关键要素,包括数据表示、缺失与整合、潜在偏差及其对结果稳健性和普适性的影响。应明确预期用途(输入类型、最终用户、自主性与人工监督水平等),并阐述低质量或缺失数据的处理方法及临床可用性。还需说明用户的交互要求与专业知识水平,并在讨论中展望未来研究方向及模型推广潜力。示例分析:例文明确了模型在临床的应用场景,可实时监测并辅助医疗报销,讨论了适用性与局限性。但未涉及模型关键要素、使用成本及稳健性,可信度不足。

3 小结

3.1 TRIPOD-LLM 意义及局限性 TRIPOD-LLM 相较于传统的 TRIPOD 报告规范,增加了对 LLMs 特有的提示工程、生成评估、公平性考量等信息的报告要求,这有助于提升 LLMs 研究的透明度

与可重复性,也为评估该类型的文章质量提供了依据。指南提供了适合不同 LLMs 类型使用的报告条目,能够指导研究者避免因归类错误导致研究信息的遗漏,以增强报告的严谨性,推动 LLMs 研究的数据标准化与模型结果的复现。但 TRIPOD-LLM 在护理领域的应用仍存在一定局限。该指南主要基于专家共识制订,尚缺乏大规模循证证据支撑,其在实际提升护理研究质量的效果仍需验证。其次,随着多模态 LLMs 在医疗领域的逐步发展,若仅使用目前的报告条目,可能导致多模态研究的关键技术细节与风险信息被遗漏,对多模态预测模型的适配性有待提升。与此同时,如何在不同国家和地区在医疗数据安全、模型资源、伦理监管规范等方面存在差异的前提下,提供开放式、透明化的技术信息仍是该指南推广与实际应用中的关键挑战。

3.2 我国护理领域应用 LLMs 开展风险预测模型相关研究的启示 近年来,我国护理领域风险预测模型研究不断深化,LLMs 在病历解析、护理信息提取及风险预测中的应用逐步展开,其引入为模型研究提供了新的技术路径,依据 TRIPOD-LLM 报告指南梳理研究流程与报告内容,有助于提高模型研究的透明度和可重复性。然而,护理研究数据通常包含患者特征与护理记录等敏感信息,因此在研究过程中必须严格遵守数据安全与伦理合规要求。在此基础上,为进一步提高研究的可追溯性与透明度,研究者应优先采用开源 LLMs,并在研究中完整披露模型版本、训练语料、参数设定及性能评估方法,以确保研究过程的可验证性与结果的科学性。综上,通过强化模型开发与报告的规范化,不仅能够减少 LLMs 性能评估中的主观偏差,还可提升研究结果的可靠性与临床可解释性,从而为护理智能化预测模型的临床转化与多场景应用奠定坚实基础。

【参考文献】

- [1] 苏龙翔,翁利,李文雄,等.大型语言模型在重症医学中的应用与挑战[J].中华医学杂志,2023,103(31):2361-2364.
- [2] 梁清钊,董广元,李楠雁,等.人工智能在肌少症精准预测与智能管理中的应用进展[J].军事护理,2025,42(1):96-99.
- [3] 陈子佳,彭文茜,张德政,等.大语言模型在中医药领域的应用、挑战与前景[J].协和医学杂志,2025,16(1):83-89.
- [4] 史纪元,罗家音,王雪莲,等.国内外人工智能护理相关研究热点和趋势分析[J].军事护理,2023,40(7):16-19.
- [5] FRASER H, CROSSLAND D, BACHER I, et al. Comparison of diagnostic and triage accuracy of ada health and webmd symptom checkers, ChatGPT, and physicians for patients in an emergency department: clinical data analysis study[J/OL]. [2025-07-03]. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10582809/>. DOI:10.2196/49995.
- [6] MENG X, YAN X, ZHANG K, et al. The application of large

language models in medicine:a scoping review[J/OL].[2025-04-23]. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11091685/>. DOI:10.1016/j.isci.2024.109713.

[7] 万君丽,卞薇.人工智能在慢性病患者营养精准管理领域中的应用进展[J].军事护理,2023,40(10):92-95.

[8] 刘泽垣,王鹏江,宋晓斌,等.大语言模型的幻觉问题研究综述[J].软件学报,2025,36(3):1152-1185.

[9] 刘泓泽,王耀国,唐圣晟,等.医学大语言模型的应用现状与发展趋势研究[J].中国数字医学,2024,19(8):1-7,13.

[10] 王耀祖,李擎,戴张杰,等.大语言模型研究现状与趋势[J].工程科学学报,2024,46(8):1411-1425.

[11] COLLINS G S, REITSMA J B, ALTMAN D G, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement[J/OL]. [2025-01-17]. <https://pubmed.ncbi.nlm.nih.gov/25569120/>. DOI:10.1136/bmj.g7594.

[12] GALLIFANT J, AFSHAR M, AMEEN S, et al. The TRIPOD-LLM reporting guideline for studies using large language models [J]. Nat Med, 2025, 31(1): 60-69.

[13] COLLINS G S, MOONS K G M, DHIMAN P, et al. TRIPOD+

(上接第 60 页)

[3] 石秀秀,丁晓彤,李峥.《中华护理教育》创刊 20 年刊文分析[J].中华护理教育,2024,21(1):10-16.

[4] SMITH S A, DUNCAN A A. Systematic and scoping reviews: a comparison and overview[J]. Semin Vasc Surg, 2022, 35(4): 464-469.

[5] TRICCO A C, LILLIE E, ZARIN W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation [J]. Ann Intern Med, 2018, 169(7): 467-473.

[6] ARKSEY H, O’Malley L. Scoping studies: towards a methodological framework[J]. Int J Soc Res Methodol, 2005, 8(1): 19-32.

[7] PETERS M D J, GODFREY C, MCINERNEY P, et al. JBI manual for evidence synthesis (2024) [EB/OL]. [2025-08-18]. <https://synthesismanual.jbi.global>.

[8] TRICCO A C, LILLIE E, ZARIN W, et al. A scoping review on the conduct and reporting of scoping reviews[J/OL]. [2025-10-07]. <https://pubmed.ncbi.nlm.nih.gov/26857112/>. DOI: 10.1186/s12874-016-0116-4.

[9] 仇如霞,顾艳莹.范围综述报告规范 (PRISMA-ScR) 的解读[J].中国循证医学杂志,2022,22(6):722-730.

[10] LEVAC D, COLOUHOUN H, O’BRIEN K K. Scoping studies: advancing the methodology [J/OL]. [2025-10-07]. <https://pubmed.ncbi.nlm.nih.gov/20854677/>. DOI: 10.1186/1748-5908-5-69.

[11] DAUDT H M L, VAN MOSSEL C, SCOTT S J. Enhancing the scoping study methodology: a large, inter-professional team’s experience with Arksey and O’Malley’s framework [J/OL]. [2025-10-07]. <https://pubmed.ncbi.nlm.nih.gov/23522333/>. DOI:10.1186/1471-2288-13-48.

[12] HARLEY J. How to undertake a scoping review[J]. Nurs Stand, 2024, 39(7): 46-49.

[13] MUNN Z, POLLOCK D, KHALIL H, et al. What are scoping re-

AI statement; updated guidance for reporting clinical prediction models that use regression or machine learning methods[J/OL]. [2025-04-16]. <https://pubmed.ncbi.nlm.nih.gov/38626948/>. DOI:10.1136/bmj-2023-078378.

[14] ZHENG N S, KELOTH V K, YOU K, et al. Detection of gastrointestinal bleeding with large language models to aid quality improvement and appropriate reimbursement [J]. Gastroenterology, 2025, 168(1): 111-120.

[15] 陈润生. 医疗大数据结合大语言模型的应用展望[J]. 四川大学学报:医学版, 2023, 54(5): 855-856.

[16] 黄河燕,李思霖,兰天伟,等.大语言模型安全性:分类、评估、归因、缓解、展望[J].智能系统学报,2025,20(1):2-32.

[17] 赵睿卓,曲紫畅,陈国英,等.大语言模型评估技术研究进展[J].数据采集与处理,2024,39(3):502-523.

[18] 施昊昊,屠馨怡,史佳伟,等.大语言模型临床实践应用范围综述[J].医学信息学杂志,2024,45(9):19-26.

[19] 吴钰,陈霞,张叶,等.国内外智慧护理研究热点及发展趋势的可视化分析[J].军事护理,2024,41(5):85-89.

(本文编辑:刘于晶)

views? Providing a formal definition of scoping reviews as a type of evidence synthesis[J]. JBI Evid Synth, 2022, 20(4): 950-952.

[14] AROMATARIS E, MUNN Z. Scoping Reviews(2017) [EB/OL]. [2025-08-18]. <https://reviewersmanual.joannabriggs.org>.

[15] PETERS M D J, MARNIE C, TRICCO A C, et al. Scoping reviews: reinforcing and advancing the methodology and application[J/OL]. [2025-10-07]. <https://pubmed.ncbi.nlm.nih.gov/34625095/>. DOI: 10.1186/s13643-021-01821-3.

[16] 曾莹,施慧,彭剑飞,等.临床研究中中医护理适宜技术安慰剂效应控制的范围综述[J].军事护理,2023,40(11):79-82,108.

[17] 庞冬,张秋雯,金三丽,等.2015—2019 年我国护理核心期刊中系统评价类研究的范围综述[J].中国护理管理,2020,20(4):496-501.

[18] MCGOWAN J, SAMPSON M, SALZWEDEL D M, et al. PRESS peer review of electronic search strategies: 2015 guideline statement[J]. J Clin Epidemiol, 2016, 75(4): 40-46.

[19] FABIANO N, GUPTA A, BHAMBRA N, et al. How to optimize the systematic review process using AI tools[J/OL]. [2025-10-07]. <https://pubmed.ncbi.nlm.nih.gov/38827982/>. DOI: 10.1002/jcv2.12234.

[20] MOZGAI S, KAURLOTO C, WINN J, et al. Machine learning for semi-automated scoping reviews[J/OL]. [2025-07-07]. <https://www.sciencedirect.com/science/article/pii/S2667305323000741?via%3Dihub>. DOI:10.1016/j.iswa.2023.200249.

[21] ANDKHOIE M, SZAFRON M, FERGUSSON D, et al. Factors underlying treatment decision-making for localized prostate cancer in the U.S. and Canada: a scoping review using principal component analysis[J]. Can Urol Assoc J, 2019, 13(7): E220-E225.

(本文编辑:刘于晶)