

中医临床辅助决策系统评价现状与趋势*

禄莎月 田茂祥 朱垚* 陆明
(南京中医药大学第一临床医学院, 江苏 南京 210029)

摘要: 目的/意义 系统梳理临床辅助决策系统 (clinical decision support system, CDSS) 的评价内容、方法, 重点分析中医临床辅助决策系统 (traditional Chinese medicine clinical decision support system, TCMCDSS) 的评价现状与发展趋势, 明确其评价重点、现存不足及中医特色评价需求, 为后续 TCMCDSS 评价研究设计、临床验证及规范化应用提供参考。方法/过程 采用文献综述与主题归纳方法, 检索中英文数据库, 从评价内容、研究方法和评价层级等方面梳理 CDSS 评价研究, 并结合代表性 TCMCDSS 研究分析其中医特色评价内容及发展趋势。结果/结论 CDSS 评价已由单一模型性能验证拓展至知识质量、系统可用性、临床决策、患者结局、实施效果、经济学及安全、伦理与监管等多维评价。TCMCDSS 评价仍以模型与知识层为主, 未来应在通用框架基础上重点关注四诊信息质量、辨证与病机判断、“症、机、法、方、药”逻辑一致性、方药安全性及临床解释性, 并加强真实世界和全生命周期评价。

关键词: 临床辅助决策系统; 中医临床辅助决策系统; 辨证论治; 评价方法; 人工智能

Evaluation of Traditional Chinese Medicine Clinical Decision Support Systems: Current Status and Future Perspectives

LU Shayue, TIAN Maoxiang, ZHU Yao, LU Ming
(The First Clinical Medical College, Nanjing University of Chinese Medicine,
Nanjing 210029, Jiangsu, China)

[Redacted text block]

* 基金项目: 国家级非物质文化遗产代表性项目[205IX-2]; 江苏省六大人才高峰项目 (RJFW-40); 江苏省 “333 高层次人才培养工程” (2018III-0121); 江苏省卫生健康委科研项目 (Z2020024); 南京市非物质文化遗产代表性项目[NJVIII-20]; 作者简介: 禄莎月 (2001-), 女, 硕士研究生, E-mail: shayue631@163.com

** 通讯作者: 朱垚, 副教授, 副主任中医师, E-mail: zhongyiyaochuanren@126.com

10.12201/bmr.202610.00046V1

[REDACTED]

1 引言

临床辅助决策系统（clinical decision support system，CDSS）是应用信息技术整合医学知识、患者信息和诊疗规则，为医务人员提供提示、预警、诊断、治疗或管理建议的计算机辅助信息系统^[1]。随着电子病历、机器学习、知识图谱和大语言模型等技术发展，CDSS 已由用药警报和规则提醒拓展至慢病管理、急诊分诊、重症监测、精准治疗和医疗质量管理等场景^[2]。

CDSS 的临床价值不能仅由算法准确率或预测性能决定。部分系统在开发或回顾性验证中表现良好，但进入真实临床环境后，仍可能因可用性不足、警报疲劳、证据更新滞后及与 workflow 脱节等问题难以形成稳定获益^{[3][4]}。因此，CDSS 评价逐步扩展至知识质量、系统可用性、临床决策、患者结局、实施效果、经济学以及安全、伦理与监管等方面。人工智能的发展又增加了知识来源可信、生成内容准确、推理可解释和性能持续稳定等评价要求。

中医临床辅助决策系统（traditional Chinese medicine clinical decision support system，TCMCDSS）近年来广泛运用于名老中医经验传承、中医专病诊疗、辨证诊断、方药推荐等领域^[5]。医临床决策以辨证论治为核心，需要综合望、闻、问、切等信息，在分析病位、病性、邪正及病势的基础上判断病机，进而确立治法、选方遣药并随证加减。因此，仅采用疾病诊断准确率、处方重合度等指标，难以全面反映 TCMCDSS 是否符合中医临床思维和诊疗规律。

目前，TCMCDSS 相关研究数量不断增加，但不同研究在评价对象、指标选择和研究设计方面存在较大差异，对辨证病机、诊疗逻辑、方药安全以及临床应用价值等内容的关注程度亦不一致。本文梳理 CDSS 主要评价内容、方法及评价层级，重点分析 TCMCDSS 评价现状和中医特色评价内容，并讨论其存在的问题与发展趋势，为后续评价研究设计、临床验证及规范化应用提供参考。

2 文献来源与分析方法

本文采用一般性文献综述与主题归纳方法。检索中国知网、万方、PubMed 和 Web of Science 数据库，中文以“临床辅助决策系统”“临床决策支持系统”“评价”“评估”“验证”“可用性”“中医”“中医药”“辨证”“证候”“方药”等为主要检索词，英文以“clinical decision support system”“evaluation”“validation”“usability”“traditional Chinese medicine”“syndrome differentiation”等为主要检索词，并结合“人工智能”“大语言模型”“知识图谱”“检索增强生成”等相关词进行补充检索。检索时间为 2021 年 1 月 1 日至 2026 年 6 月 10 日，同时通过参考文献追溯补充 CDSS 评价相关经典方法学文献。

文献选择以能够反映 CDSS 评价内容、方法、层级及临床应用价值的代表性原始研究、综述和方法学文献为主；TCMCDSS 部分重点关注辨证诊断、方药推荐、中医知识库、智能问诊及临床决策评

价等研究。对于缺乏明确临床决策支持场景、仅介绍系统功能而无评价内容或与本文主题关联较弱的文献，不作为重点分析对象。采用描述性归纳与主题分析方法总结相关研究。

3 CDSS 评价内容、方法及评价层级

3.1 CDSS 主要评价内容及常用方法

3.1.1 技术性能评价

技术性能评价是 CDSS 进入临床应用前的基础环节，主要判断系统能否稳定、准确地完成风险预测、诊断识别、分层预警或治疗推荐等任务。常用指标包括准确率、灵敏度、特异度、受试者工作特征曲线下面积（area under the receiver operating characteristic curve，AUC）、F1 值、校准度和决策曲线分析等^{[6][7]}。现有研究多采用回顾性数据库、独立外部验证队列或多中心数据进行评价。Yang 等^[8]基于中国 7 家三级医院心力衰竭住院患者数据开发静脉用药风险分层 CDSS，人工神经网络模型在外部验证中的 AUC 达 0.921。Zhang 等^[9]利用多个重症监护数据库构建院内死亡风险预测模型，并采用 SHAP 方法解释关键预测变量；Wei 等^[10]进一步将评价对象拓展至大语言模型，比较其急诊内科疾病诊断准确性、鉴别诊断完整性和响应时间；可见，技术性能评价能够为 CDSS 的模型选择、系统优化和临床转化提供前期依据，正由单纯比较模型区分能力，逐步向性能、泛化、校准和可解释性并重转变。

3.1.2 知识库与规则质量评价

知识库与规则质量评价是知识驱动型 CDSS 产生临床建议的重要基础，其评价重点在于知识来源是否权威、内容是否完整、规则逻辑是否合理以及证据是否可追溯和持续更新。常用方法包括指南对照、专家评审、内容效度评价、案例测试和规则审计等。

Kuria 等^[11]在肯尼亚新生儿照护场景中评价人类监督的大语言模型 CDSS，将系统建议与国家方案进行对照，并由专家评审正确性和引用准确性；Baklaci 等^[12]将大语言模型生成的头颈部肿瘤 TNM 分期和治疗建议与 AJCC 分期标准、NCCN 指南及多学科决策比较，显示专科 CDSS 需进行疾病亚型和场景化验证。Felde 等^[13]以风湿病指南构建检索增强生成（retrieval-augmented generation，RAG）知识库，并采用 F1 值和 RAGAS 评分评价检索质量与回答适宜性。随着大语言模型与 RAG 技术进入 CDSS，知识质量评价还需纳入检索相关性、引用准确性和生成内容一致性。

3.1.3 系统可用性评价

系统可用性评价关注 CDSS 是否易学、易用并能够顺畅融入临床工作流程，是连接技术性能与临床采纳的重要环节。常用方法包括系统可用性量表（system usability scale，SUS）、NASA-TLX 认知负荷量表、任务完成率、操作时间、点击或浏览次数、think-aloud 测试、满意度调查及半结构式访谈等。

Larsen 等^[14]在 2 型糖尿病用药 CDSS App 开发中采用用户中心设计，通过多轮测试比较指南依从性、任务时间、SUS 评分和认知负荷。Köhler 等^[15]利用远程 think-aloud、访谈和 SUS 评价基层医疗 CDSS 原型，识别术语表达、数据录入和导航设计等问题。Wilson-Jene 等^[16]通过临床医生访谈和问卷开展迭代优化。由此可见，面向生成式和多模态 CDSS，还应进一步关注解释可理解性、信任形成机制、认知负荷、工作流适配和医患共同决策体验，以提高系统在真实临床环境中的可接受性和可持续应用能力。

3.1.4 临床决策过程评价

临床决策过程评价关注 CDSS 是否改变或优化医务人员诊疗行为，是连接系统性能与患者结局的重要中间环节。常用指标包括诊断正确率、鉴别诊断完整性、治疗方案合理性、指南依从性、处方合理性、医生决策时间、建议采纳率、决策一致性等。

Zamora 等^[17]通过模拟病例前后对照评价血脂管理 CDSS，发现系统可降低专科医生间处方差异。Jeong 等^[18]系统评价随机对照试验发现，人工智能临床辅助决策系统（artificial intelligence-clinical decision support system, AI-CDSS）对医务人员诊断准确性具有一定提升作用，但效应量较小且异质性较高。高博等^[19]将临床决策辅助系统用于脊柱外科规培教学，显示其可改善临床推理训练和规范化接诊。该类评价更强调 CDSS 在具体临床任务中对医生判断、诊疗路径和处方行为的实际影响，应重点考察 CDSS 是否真正改善医生决策质量、减少决策变异并提高指南执行水平，但其结果仍需进一步结合临床结局和真实世界实施效果综合判断。

3.1.5 患者结局评价

患者结局评价关注 CDSS 是否最终改善健康结局和医疗服务结局，是判断系统临床价值的关键。常用指标包括症状改善、疾病控制率、实验室指标变化、住院时间、ICU 转入率、再入院率、不良事件、用药错误、生活质量和患者报告结局等。

Gerber 等^[20]在农村 2 型糖尿病管理中采用整群随机试验，以 12 个月糖化血红蛋白为主要结局评价社区卫生工作者使用 CDSS 后的照护效果。Aberer 等^[21]以血糖达标率、平均血糖和低血糖发生率评价住院患者的胰岛素 CDSS 疗效与安全性。Shakarbaev^[22]通过随机对照试验显示 AI-CDSS 可降低用药错误和药物不良事件，缩短住院时间并降低 30 天再入院率。总体而言，临床结局评价更强调 CDSS 对患者获益和医疗安全的实际影响，应优先采用前瞻性、随机对照或真实世界研究设计，并结合安全性、患者体验和长期随访指标，以避免仅凭模型性能或过程改善推断临床获益。

3.1.6 实施效果评价

实施效果评价关注 CDSS 在真实医疗环境中能否被稳定部署、持续使用并融入临床流程。常用指标包括系统使用频率、医生接受度、建议采纳率、工作流匹配度、跨机构推广性、持续使用率和警报疲劳等。

Feldman 等^[23]通过持续监测新获批靶向治疗证据并主动通知医生，体现 CDSS 在证据更新中的实施价值。Lim 等^[24]开发的居家医院患者识别 CDSS，实施后转诊数量、转诊及时性和服务利用度均有改善。Newton 等^[25]指出，系统实用性、易用性、电子病历集成度和工作流匹配度是影响医生长期接受与使用的核心因素。因此，实施评价应贯穿部署前、中、后全过程，并结合组织环境、人员培训和持续反馈综合判断系统的可推广性。

3.1.7 经济学评价

经济学评价关注 CDSS 在改善诊疗质量和患者结局的同时，是否具有合理资源投入和经济价值。常用指标包括系统开发成本、实施成本、维护成本、人员培训成本、资源消耗、门诊或住院费用、预算影响和投资回报等。

White 等^[26]综述指出，医院 CDSS 经济学评价多从医院视角分析成本与后果，但成本构成、资源估值、时间范围、结局选择和不确定性分析差异较大。Donovan 等^[27]进一步发现，人员劳动、培训、交互式支持和系统迭代是医院 CDSS 实施成本的重要来源；未来需进一步规范 CDSS 经济学评价的方法学，可参照 CHEERS 2022 报告规范^[28]，将开发、部署、维护、更新乃至去实施等全生命周期成本纳入评价，以提高经济学研究的规范性和可比性。

3.1.8 安全、伦理与监管评价

安全、伦理与监管评价关注 CDSS 在真实应用中可能产生的误导风险、患者伤害、责任归属、数据安全、隐私保护、算法偏倚、可解释性不足和上市后监测等问题，是保障系统可信、可控和可持续应用的重要维度。

Lyell 等^[29]基于 FDA MAUDE 数据库分析机器学习医疗器械安全事件，提示风险不仅来源于算法本身，还与数据输入、设备集成、用户交互和临床流程相关。Elgin 等^[30]从医务人员视角指出，AI-CDSS 参与资源分配可能引发公平性、知情同意、隐私保护、责任归属和黑箱决策等伦理问题。Baliga 等^[31]认为，麻醉 AI-CDSS 从预测模型走向床旁应用时，常受泛化能力不足、前瞻性证据缺乏、解释性不足和监管责任不明等限制。因此，CDSS 安全与监管评价应覆盖开发、部署、监测和更新阶段，并结合技术审查、伦理审查、真实世界监测和责任治理框架。

CDSS 临床应用评价方法分类总结见表 1。

表 1 CDSS 主要评价内容、方法与指标

评价类型	核心问题	常用方法与指标
技术性能评价	模型或算法是否可靠	准确率、灵敏度、特异度、AUC、F1 值、校准度、外部验证、决策曲线分析
知识库与规则质量评价	知识来源、规则和证据是否可信	指南一致性、专家一致性、内容效度、规则完整性、更新机制、可追溯性
系统可用性评价	医生是否愿意用、能否顺畅用	SUS、NASA-TLX、任务完成率、操作时间、think-aloud、访谈、满意度
临床决策过程评价	是否改善医生诊疗行为	诊断正确率、处方合理率、指南依从性、警报采纳率、决策时间
患者结局评价	是否使患者获益	症状评分、再入院率、死亡率、不良事件、生活质量、患者报告结局
实施效果评价	是否可持续落地	采纳率、使用频率、 workflow 整合、警报疲劳、跨机构推广性、实施日志
经济学评价	投入是否具有价值	成本分析、成本-效果、成本-效用、预算影响、投资回报
安全伦理与监管评价	是否安全、合规、可追责	数据安全、隐私、偏倚、可解释性、责任归属、上市后监测

3.2 CDSS 评价层级

上述 8 类评价内容分别回答了 CDSS 在性能、知识、使用、临床价值和持续运行等方面的不同问题，但各类评价并非彼此孤立，而是伴随系统由研发走向临床应用呈现明显的层级递进关系。综合现有研究，可将 CDSS 评价概括为模型/知识层、系统层、医生决策层、患者结局层、医疗系统层和全生命周期层 6 个相互衔接的评价层级。CDSS 不同评价层级及核心内容见表 2。

表 2 CDSS 不同评价层级及核心内容

评价层级	核心目标	核心评价内容	常用研究方法	参考文献
模型/知识层	验证算法与知识库的科学可靠性	算法准确率与泛化性、指南一致性、知识完整性、数据质量	回顾性验证、内部/外部验证、多中心验证、交叉验证、专家评审、指南一致性评价、内容效度验证、偏倚风险评价、报告规范评价	[6][7][8][11][32][33]
系统层	保障系统易用、稳定、可集成	界面易用性、任务效率、系统稳定性、电子病历集成度、认知负荷	用户中心设计、原型测试、可用性测试、半结构式访谈、焦点小组、任务分析、SUS 量表、认知负荷评价、HOT-fit 评价	[14][15][34][35]
医生决策层	改善诊疗行为与决策质量	诊断正确率、指南依从性、处方合理性、决策效率、警报采纳率	模拟病例研究、前后对照研究、随机对照试验、交叉试验、诊疗方案一致性评价、医生行为观察、系统日志分析、Meta 分析	[17][18][36][37][38]
患者结局层	改善患者健康结局与体验	疾病控制率、不良事件发生率、住院时间、生活质量、患者满意度	个体随机对照试验、整群随机试验、队列研究、真实世界研究、患者报告结局研究	[20][21][22][39]
医疗系统层	提升医疗体系质量与经济性	系统采纳率、工作流适配度、成本效果比、资源利用效率	实施研究、混合方法研究、成本-效果分析、成本-效用分析、成本-后果分析、卫生经济学评价、采纳与持续使用研究	[25][26][27][28][40][41]
全生命周期层	保障持续安全有效运行	更新机制、安全监测、算法漂移、伦理合规、上市后再评价	上市前-上市后连续评价、真实世界研究、上市后监测、安全事件分析、伦理质性研究、风险治理框	[29][30][32][42][43]

10.12201/bmr.202610.00046V1

4 TCMCDSS 评价研究现状及特色评价内容

4.1 TCMCDSS 评价研究现状

现有 TCMCDSS 评价主要集中在证候识别、方药推荐、推理与知识可信性以及临床决策等方面。较早研究多以准确率、召回率、F1 值和专家一致性评价疾病或证候识别。Zhang 等^[44]基于中医病历和临床指南构建疾病诊断模型，并通过不同算法比较评价预测性能；Li 等^[45]将中医知识融入循环神经网络开展脑性瘫痪证候诊断。Long 等^[46]构建银屑病辨证原型系统，在算法之外增加专家对知识体系、推理规则、界面和响应速度的评价。这些研究主要验证输入与诊断结果的对应关系，对病机形成、治法选择和动态辨证关注较少。

处方推荐研究开始将中间诊疗环节纳入模型。Tian 等^[47]构建 TCMLLM-PR，采用 Precision、Recall 及 F1 等指标评价 LLM 处方推荐及跨数据集迁移能力；Yue 等^[48]提出 SDPR，将中医诊疗过程表征为症状、证候、治法和中药之间的递进关系，并在公共及真实世界数据上评价处方推荐性能。但此类研究的评价终点仍多为推荐药物与参考处方的一致程度，对治法、剂量和随证加减合理性的评价相对不足。

随着 LLM 和 RAG 用于中医诊疗，评价开始关注知识来源、推理过程和解释能力。Jia 等^[49]构建 Qibo 及相应评价基准，从主观评价、客观知识和中医自然语言处理任务等方面测试模型；Wang 等^[50]构建 TCMEval-SDT，将辨证过程细分为临床信息提取、病机推理、证候推理和解释性总结。Li 等^[51]进一步将知识图谱、RAG 和思维链应用于个体化辨证推理，都提示生成式 TCMCDSS 除结果准确性外，还需评价知识依据和推理过程。

少数研究开始进入临床决策和人机协同层面。刘奕清等^[52]构建基于 LLM 的中医智能预问诊系统，除 BLEU-4、ROUGE-L 和专业知识测试外，还通过临床测试比较系统与主治医师的辨证准确率、问诊轮次和问诊时间。Liu 等^[53]进一步比较 5 种 LLM 及医师—AI 协作模式，将临床决策质量、决策时间、SUS 以及建议采纳、修改和拒绝率纳入评价，提示部分研究已开始评价 TCMCDSS 对医师实际决策行为的影响。姚斌莲等^[54]对中医护理决策支持系统进行实施前后比较，并观察护理效果、方案满意度、病历书写时间和护士体验，是较少涉及真实临床应用效果的研究之一。

总体而言，TCMCDSS 评价仍以模型和知识层为主，医生决策、患者结局、实施及长期应用评价相对不足。现有研究虽已开始关注推理过程和人机协作，但能否改善中医临床决策、提高诊疗效率并带来患者获益，仍缺乏充分验证。

4.2 TCMCDSS 特色评价内容

中医诊疗以辨证论治为核心，其信息来源、决策过程和治疗选择具有自身特点。TCMCDSS 评价不宜脱离通用 CDSS 评价内容另设体系，而应在模型性能、知识质量、可用性和临床效果等评价基础上，进一步关注辨证与病机判断、四诊信息质量、诊疗逻辑一致性与临床解释性，以及方药安全性和知识可信性。

4.2.1 以辨证论治为核心的综合决策

中医诊疗不以疾病名称作为唯一决策依据，而是在四诊合参基础上辨析病位、病性、病势和邪正关系，形成证候诊断，并进一步确立治则治法和方药。系统输出不应仅包括疾病诊断，还应包含证候类型、证候要素、病机分析、治法原则、方药建议及加减依据等多层内容^[55]。与西医 CDSS 常见单标签疾病预测不同，证候具有动态演变特征，系统需支持同病异治、异病同治和随证加减。

4.2.2 以四诊合参为基础的多模态输入

TCMCDSS 需处理舌象、面色、声音、脉象、症状和病史等图像、声学信号、生理波形及文本信息。四诊信息存在主观性较强、术语不统一和采集质量差异等问题^[56]。评价时应关注数据采集规范、术语标准、缺失与噪声控制以及不同模态信息的一致性。对于多模态系统，还需评价舌诊、脉诊、问诊等信息融合后是否确实提高辨证和决策质量^[57]，并关注不同模态信息缺失、冲突或质量下降时系统输出的稳定性。

4.2.3 以理法方药贯通为核心的可解释推理

TCMCDSS 的治疗推荐不能仅表现为“给出某方”或“推荐某药”，还应说明所依据的病机、治法、方证关系、君臣佐使、药性功效和加减逻辑。知识图谱、规则引擎和大语言模型等技术为理法方药关系建模提供了工具，但如果缺乏可追溯证据和专家审查，可能出现辨证与处方脱节、方药与治法不符、加减依据不清等风险^[58]。因此，TCMCDSS 评价应重点关注诊疗逻辑一致性、临床解释性及决策依据的可追溯性。

4.2.4 以方药安全与知识传承为导向的价值评价

中医治疗强调因人、因时、因地制宜，TCMCDSS 安全评价不仅要关注一般药物不良反应，还应关注证候适配性、配伍禁忌、剂量合理性、加减是否得当及长期用药风险。同时，TCMCDSS 还承担名老中医经验、临床医案和古籍知识数字化转化功能^[59]。其知识来源往往包括现代指南、专家共识、真实世界病案、名医经验、经典古籍和中药复方人用经验，系统评价应兼顾理论依据、人用经验和临床验证证据，形成完整评价闭环。

5 TCMCDSS 评价面临的问题与未来展望

目前 TCMCDSS 评价仍存在临床证据不足和评价标准不统一等问题。多数研究集中于模型准确性、证候识别和方药推荐，针对医生决策改善、患者结局、医疗资源利用和长期应用效果的研究较少，多中心、前瞻性和真实世界证据不足。不同研究在病例选择、参考标准、专家评价方式和指标设置方面差异较大，四诊信息标准化、病机标注以及“症、机、法、方、药”一致性的评价方法尚未统一，影响研究结果比较。

人工智能技术进一步增加了评价难度。当前中医大模型研究仍以文本和静态病例为主，多模态四诊数据及连续诊疗场景应用不足^[60]。RAG 和生成式模型应用后，除最终答案外，还需评价知识来源、检索结果和生成内容之间的一致性与可追溯性。症状提取或病机判断的早期偏差还可能影响后续治法、方药、剂量及禁忌判断。已有研究尝试通过多轮交互和领域知识增强提高临床适配性^[61]，但生成稳定性、不确定性表达、高风险建议处理及模型更新后的持续有效性仍需进一步研究。

未来应加强标准化评价基础，建立覆盖四诊信息、病机判断、治法选择和方药应用的高质量病例资料，并通过多专家讨论形成合理的参考标准。评价内容应与系统应用阶段相匹配：研发阶段重点考察模型性能、知识质量和诊疗逻辑；临床应用阶段重点考察可用性、医生决策、人机协同和安全性；推广应用后进一步评价患者获益、资源利用、经济价值及长期安全。对于持续更新的人工智能系统，还应记录模型和知识库版本，并结合真实世界数据开展持续评价。通过多中心、前瞻性临床验证和真实世界研究，逐步积累从技术性能到临床应用和长期安全的评价证据，并避免以单一自动指标或静态题库结果替代真实临床价值判断。

需要指出的是，本文基于现有文献和中医诊疗特点归纳 TCMCDSS 评价重点，旨在总结研究现状并提出未来方向，尚未通过专家咨询、指标赋权或临床验证形成正式评价指标体系，相关内容仍需进一步验证。

6 结语

TCMCDSS 评价正在由单一模型性能验证向多维度、多层级和全过程评价转变。未来应在现有 CDSS 评价内容和方法的基础上，结合中医辨证论治、四诊合参、“症、机、法、方、药”逻辑及方药安全等特点，进一步加强临床决策、患者结局、真实世界应用和全生命周期评价，逐步形成与系统成熟阶段相匹配的评价证据链，从而推动 TCMCDSS 由技术可行性验证向规范、安全且具有真实临床价值的应用转化。

参考文献

[1] 国家卫生健康委办公厅. 医疗机构临床决策支持系统应用管理规范（试行）[S/OL]. (2023-07-17)[2026-06-14]. <https://www.nhc.gov.cn/yzygj/c100068/202307/d5239ed63d324aaa8ea85922b5fb1310.shtml>

[2] 刘伊,王文锐.我国临床决策支持系统的应用热点分析与发展趋势——基于 CiteSpace 的可视化分析[J].中国数字医学,2026,21(4):98-105.

[3] LEE T C, SHAH N U, HAACK A, et al. Clinical implementation of predictive models embedded within electronic health record systems: a systematic review[J]. Informatics, 2020, 7(3): 25.

[4] BLECKER S, PANDYA R, STORK S, et al. Interruptive versus noninterruptive clinical decision support: usability study[J]. JMIR Human Factors, 2019, 6(2): e12469.

[5] 郑婉婷,李敬华,田少磊,等. 中医临床辅助决策系统功能与算法的发展评述[J]. 中国数字医学,2022,17(2):86-94.

[6] COLLINS G S, MOONS K G M, DHIMAN P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods[J]. BMJ, 2024, 385: e078378.

[7] MOONS K G M, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods[J]. BMJ, 2025, 388: e082505.

[8] YANG Y, XU Z J, YU P L, et al. Development and validation of a machine learning-based clinical decision support tool for stratifying intravenous medication risk in hospitalized patients with heart failure [J]. International Journal of Clinical Pharmacy, 2026.

[9] ZHANG T, NAN Z, GUO S, et al. Multicenter machine learning model for assessing the impact of malignancy on in-hospital mortality in heart failure patients: a clinical decision support system with interpretable artificial intelligence [J]. Scientific Reports, 2026, 16 (1): 5234.

[10]WEI J T, JIANG S Y, YIN T, et al. Performance evaluation of large language models in the diagnosis of emergency internal medicine diseases: a retrospective study [J]. Frontiers in Public Health, 2026, 14: 1780425.

[11]KURIA T, KAMAU G, MAKOKHA F, et al. Human-supervised, large language model-based clinical decision support aligned to national newborn protocols in Kenya: a pragmatic, early-stage evaluation [J]. Frontiers in Digital Health, 2026, 8: 1832634.

[12]BAKLACI D, IŞIK H, ERDEM D, et al. Accuracy of large language models in head and neck cancers: a comparative analysis of ChatGPT and Gemini in TNM staging and clinical decision support [J]. Frontiers in Oncology, 2026, 16: 1828538.

[13]FELDE S, BUCHKREMER R, CHEHAB G, et al. Low-energy small language models with retrieval-augmented generation can surpass large-model performance in rheumatology [J]. Frontiers in Medicine, 2026, 13: 1817215.

[14]LARSEN K, AKINDELE B, HEAD H, et al. Developing a user-centered digital clinical decision support app for evidence-based medication recommendations for type 2 diabetes mellitus: prototype user testing and validation study [J]. JMIR Human Factors, 2022, 9 (1): e33470.

[15]KÖHLER S M, HARTLEB S, SIEBEN D, et al. Evaluating the prototype of a clinical decision support system in primary care: qualitative study [J]. JMIR Formative Research, 2025, 9: e69875.

[16]WILSON-JENE H, COWAN R E, ROY Z, et al. Development of the RightWheel manual wheelchair wheel rolling resistance clinical decision support system through iterative clinician interviews [J]. Journal of Rehabilitation and Assistive Technologies Engineering, 2025, 12: 20556683251336184.

[17]ZAMORA L, MUÑOZ A, GARCÍA-RÍO F, et al. Pilot study evaluating a clinical decision support tool: "Virtua Lipidologist HTE-DLP 3.0"[J]. Cardiology Research, 2024, 15 (2): 89-97.

[18]JEONG M A, KIM S D. Impact of AI-based clinical decision support systems on diagnostic accuracy among healthcare

- professionals: a systematic review and meta-analysis of randomized controlled trials [J]. *Applied Sciences*, 2026, 16 (10): 5146.
- [19]高博,吕亚琳,丁子毅,等.基于自主研发临床决策辅助教学系统在脊柱外科规培教学中的应用评价 [J]. *骨科*.2026, 17 (3): 263-268.
- [20]GERBER F, GUPTA R, SANCHEZ-SAMANIEGO G, et al. Community health worker-led versus facility-based type 2 diabetes care in rural Lesotho: a cluster-randomized trial within the ComBaCaL cohort study [J]. *BMC Medicine*, 2026, 24 (1): 128.
- [21]ABERER F, HOCHFELLNER D A, BAUMANN P M, et al. Glycemic efficacy and safety by using insulin degludec and aspart guided by a clinical decision support system in non-critically ill inpatients with type 2 diabetes mellitus [J]. *Biosensors*, 2026, 16 (5): 289.
- [22]SHAKARBAEV N. AI-enhanced clinical decision support reduces medication errors and adverse drug events in a multicenter teaching hospital network: a prospective randomized controlled trial [J]. *International Journal of Medical Informatics*, 2026, 215: 106457.
- [23]FELDMAN R, ABDULLA F, HINTON A, et al. Proactive identification of candidates for newly approved targeted therapies using a clinical decision support program [J]. *The Oncologist*, 2026, 31 (6): oyag169.
- [24] LIM S M, KITT-THOMPSON T, JOHNSON D F, et al. Development and implementation of a computerised clinical decision support system for hospital-in-the-home patient identification [J]. *Internal Medicine Journal*, 2026.
- [25]NEWTON N, BAMGBOJE-AYODELE A, FORSYTH R, et al. A systematic review of clinicians' acceptance and use of clinical decision support systems over time [J]. *npj Digital Medicine*, 2025, 8 (1): 309.
- [26]WHITE N M, CARTER H E, KULARATNA S, et al. Evaluating the costs and consequences of computerized clinical decision support systems in hospitals: a scoping review and recommendations for future practice [J]. *Journal of the American Medical Informatics Association*, 2023, 30 (6): 1205-1218.
- [27]DONOVAN T, ABELL B, FERNANDO M, et al. Implementation costs of hospital-based computerised decision support systems: a systematic review [J]. *Implementation Science*, 2023, 18 (1): 7.
- [28]HUSEREAU D, DRUMMOND M, AUGUSTOVSKI F, et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations[J]. *Value in Health*, 2022, 25(1): 3-9.
- [29]LYELL D, WANG Y, COIERA E, et al. More than algorithms: an analysis of safety events involving ML-enabled medical devices reported to the FDA [J]. *Journal of the American Medical Informatics Association*, 2023, 30 (7): 1227-1236.
- [30]ELGIN C Y, ELGIN C. Ethical implications of AI-driven clinical decision support systems on healthcare resource allocation: a qualitative study of healthcare professionals' perspectives [J]. *BMC Medical Ethics*, 2024, 25 (1): 148.
- [31] BALIGA J, SESHADRI N. From prediction to practice: closing the translation gap in artificial intelligence for anesthesia [J]. *Journal of Clinical Monitoring and Computing*, 2026.
- [32]INTERNATIONAL MEDICAL DEVICE REGULATORS FORUM. Software as a Medical Device: clinical evaluation [S/OL]. 2017 [2026-05-20]. <https://www.imdrf.org/documents/software-medical-device-clinical-evaluation>.
- [33]WOLFF R F, MOONS K G M, RILEY R D, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies[J]. *Annals of Internal Medicine*, 2019, 170(1): 51-58.
- [34]YUSOF M M, KULJIS J, PAPAFAEIROPOULOU A, et al. An evaluation framework for health information systems: human, organization and technology-fit factors (HOT-fit)[J]. *International Journal of Medical Informatics*, 2008, 77(6): 386-398.
- [35]ÇUBUKÇU H C, CANBAZ H. Usability assessments in clinical decision support systems[J]. *Northern Clinics of Istanbul*, 2025, 12(3): 372-377.
- [36]KAWAMOTO K, HOULIHAN C A, BALAS E A, et al. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success[J]. *BMJ*, 2005, 330(7494): 765-768.
- [37]GARG A X, ADHIKARI N K J, MCDONALD H, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review[J]. *JAMA*, 2005, 293(10): 1223-1238.
- [38]JASPERS M W M, SMEULERS M, VERMEULEN H, et al. Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings[J]. *Journal of the American Medical*

Informatics Association, 2011, 18(3): 327-334.

[39]EL ASMAR M L, DHARMAYAT K I, VALLEJO-VAZ A J, et al. Effect of computerised, knowledge-based, clinical decision support systems on patient-reported and clinical outcomes of patients with chronic disease managed in primary care settings: a systematic review[J]. BMJ Open, 2021, 11(12): e054659.

[40]JIA P, ZHANG L, CHEN J, et al. The effects of clinical decision support systems on medication safety: an overview[J]. PLoS One, 2016, 11(12): e0167683.

[41]JACOB V, THOTA A B, CHATTOPADHYAY S K, et al. Cost and economic benefit of clinical decision support systems for cardiovascular disease prevention: a Community Guide systematic review[J]. Journal of the American Medical Informatics Association, 2017, 24(3): 669-676.

[42]FOOD and Drug Administration. Marketing submission recommendations for a predetermined change control plan for artificial intelligence-enabled device software functions: guidance for industry and FDA staff[S/OL]. Silver Spring: FDA, 2025[2026-05-20].

<https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-intelligence-enabled>.

[43]LABKOFF S, OLSEN L, SHAH N, et al. Toward a responsible future: recommendations for AI-enabled clinical decision support [J]. Journal of the American Medical Informatics Association, 2024, 31 (11): 2730-2739.

[44]ZHANG Z, LI J, ZHENG W, et al. Research on diagnosis prediction of traditional Chinese medicine diseases based on improved Bayesian combination model[J]. Evidence-Based Complementary and Alternative Medicine, 2021, 2021: 5513748.

[45]LI D, QU J, TIAN Z, et al. Knowledge-based recurrent neural network for TCM cerebral palsy diagnosis[J]. Evidence-Based Complementary and Alternative Medicine, 2022, 2022: 7708376.

[46]LONG H, WANG Z, CUI Y, et al. A prototype for diagnosis of psoriasis in traditional Chinese medicine[J]. Computers, Materials & Continua, 2022, 73(3): 5197-5217.

[47]TIAN H, YANG K, DONG X, et al. TCMLLM-PR: evaluation of large language models for prescription recommendation in traditional Chinese medicine[J]. Digital Chinese Medicine, 2024, 7(4): 343-355.

[48]YUE W, JI W, WANG X, et al. SDPR: prescription recommendation with syndrome differentiation in traditional Chinese medicine[J]. IEEE Journal of Biomedical and Health Informatics, 2025, 29(5): 3736-3749.

[49]JIA Y, JI X, WANG X, et al. Qibo: a large language model for traditional Chinese medicine[J]. Expert Systems with Applications, 2025, 284: 127672.

[50]WANG Z, HAO M, PENG S, et al. TCMEval-SDT: a benchmark dataset for syndrome differentiation thought of traditional Chinese medicine[J]. Scientific Data, 2025, 12: 437.

[51]LI J, CHANG Y, LIU Y, et al. TCM-DiffRAG: personalized syndrome differentiation reasoning method for traditional Chinese medicine based on knowledge graph and chain of thought[J]. Frontiers in Medicine, 2026, 13: 1804478.

[52]刘奕清,李迎,杨洪军,等.基于大语言模型的中医智能预问诊系统的构建与验证[J].中医杂志,2025,66(9):895-900.

[53]LIU W, MIAO S, MA Q, et al. Digitally assisted clinical decision-making in traditional Chinese medicine: comparative study of 5 large language models[J]. JMIR Formative Research, 2026, 10: e80167.

[54]姚斌莲,徐敏,张翔,等.中医护理决策支持系统的构建与应用[J].中国护理管理,2025,25(11):1662-1667.

[55]宋逸杰,马素亚,戴亚盛,等.人工智能辅助中医辨证的关键问题与技术挑战[J].中国工程科学,2024,26(2):234-244.

[56]杨涛,朱学芳.中医辨证智能化研究现状及发展趋势[J].南京中医药大学学报,2021,37(4):597-601.

[57]崔骥,许家佗.人工智能信息技术在中医四诊现代化研究中的应用现状与展望[J].上海中医药杂志,2025,59(1):7-12.

[58]王松,李正钧,杨涛,等.中医药知识图谱研究现状及发展趋势[J].南京中医药大学学报,2022,38(3):272-278.

[59]杨涛,任海燕,周作建,等.人工智能赋能中医学高质量发展面临的问题与挑战[J].南京中医药大学学报,2024,40(12):1285-1290.

[60]HAN C, YANG G, LI H, et al. Tuning and clinical application of large language models in Traditional Chinese Medicine: scoping review[J]. Chinese Medicine, 2026, 21: 71.

[61]WANG L, TANG K, YANG Z, et al. GastroTCM: a large language model assistant for gastroenterology in traditional Chinese medicine[J]. Chinese Medicine, 2026, 21: 46.

