

## 人工智能如何改变指南的未来

陈耀龙<sup>1,2,3,4,5</sup>, 罗旭飞<sup>1</sup>, 史乾灵<sup>6</sup>, 吕萌<sup>1</sup>, 周奇<sup>6</sup>, 王健健<sup>1</sup>, 杨楠<sup>3</sup>,  
高东平<sup>7</sup>, 杨书<sup>8</sup>, 商洪才<sup>9</sup>, 杨克虎<sup>3,4,5</sup>

<sup>1</sup> 兰州大学公共卫生学院, 兰州 730000

<sup>2</sup> 兰州大学健康数据科学研究院, 兰州 730000

<sup>3</sup> 兰州大学基础医学院循证医学中心, 兰州 730000

<sup>4</sup> 世界卫生组织指南实施与知识转化合作中心, 兰州 730000

<sup>5</sup> Cochrane 中国协作网兰州大学分中心, 兰州 730000

<sup>6</sup> 兰州大学第一临床医学院, 兰州 730000

<sup>7</sup> 中国医学科学院 北京协和医学院医学信息研究所/图书馆, 北京 100020

<sup>8</sup> 成都中医药大学, 成都 610075

<sup>9</sup> 北京中医药大学东直门医院中医内科学 教育部和北京市重点实验室, 北京 100700

通信作者: 陈耀龙 电话: 0931-8912639, E-mail: chenyaolong@vip.163.com

**【摘要】**近年来, 人工智能的快速发展和进步对医学领域的影响日益加深。科技部发布的人工智能指南明确提出要加大人工智能技术在医药卫生领域应用的支持力度。临床实践指南作为医疗实践的重要指导性文件, 可以规范医务人员的诊疗行为, 缩小最佳研究证据与当前实践之间的差距, 具有重要意义和价值。如何利用人工智能技术加速指南的制订流程, 提高其制订效率, 创新其传播与实施模式, 乃至改变指南的未来发展, 已引起国内外学者的广泛关注。本文就人工智能在临床实践指南领域的应用现状和发展前景进行分析和讨论, 并在此基础上提出促进人工智能与临床实践指南结合的思考与建议。

**【关键词】** 临床实践指南; 人工智能; 机器学习; 大数据

**【中图分类号】** R181.2; TP18 **【文献标志码】** A **【文章编号】** 1674-9081(2021)01-0114-08

**DOI:** 10.12290/xhyxzz.2021-0012

## How Will Artificial Intelligence Lead the Future of Clinical Practice Guidelines

CHEN Yao-long<sup>1,2,3,4,5</sup>, LUO Xu-fei<sup>1</sup>, SHI Qian-ling<sup>6</sup>, LYU Meng<sup>1</sup>, ZHOU Qi<sup>6</sup>, WANG Jian-jian<sup>1</sup>,  
YANG Nan<sup>3</sup>, GAO Dong-ping<sup>7</sup>, YANG Shu<sup>8</sup>, SHANG Hong-cai<sup>9</sup>, YANG Ke-hu<sup>3,4,5</sup>

<sup>1</sup> School of Public Health, Lanzhou University, Lanzhou 730000, China

<sup>2</sup> Lanzhou University Institute of Health Data Science, Lanzhou 730000, China

<sup>3</sup> Evidence-based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou 730000, China

<sup>4</sup> WHO Collaborating Center for Guideline Implementation and Knowledge Translation, Lanzhou 730000, China

<sup>5</sup> Lanzhou University, an Affiliate of the Cochrane China Network, Lanzhou 730000, China

<sup>6</sup> The First Clinical Medical College, Lanzhou University, Lanzhou 730000, China

<sup>7</sup> Institute of Medical Information & Library, Chinese Academy of Medical Sciences & Peking Union Medical College,

基金项目: 国家重点研发计划 (2018YFC1705500)

陈耀龙、罗旭飞对本文同等贡献

引用本文: 陈耀龙, 罗旭飞, 史乾灵, 等. 人工智能如何改变指南的未来 [J]. 协和医学杂志, 2021, 12 (1): 114-121. doi: 10.12290/xhyxzz.2021-0012.

Beijing 100020, China

<sup>8</sup>Chengdu University of Traditional Chinese Medicine, Chengdu 610075, China

<sup>9</sup>Key Laboratory of Chinese Internal Medicine of Ministry of Education, Beijing Dongzhimen Hospital,  
Beijing University of Chinese Medicine, Beijing 100700, China

Corresponding author: CHEN Yao-long Tel: 86-931-8912639, E-mail: chenyaolong@vip.163.com

**【Abstract】** In the past decades, with the rapid development and progress of artificial intelligence (AI), its impact on the field of medicine is deepening. The AI guideline released by the Ministry of Science and Technology of the People's Republic of China also suggested increasing support for the application of AI in medicine. As an important guiding document in medical practice, clinical practice guidelines are of great significance and value in regulating the behavior of healthcare professionals and bridging the gap between the best evidence of research and current practice. How to use AI to accelerate the development process, improve efficiency, innovate the format of dissemination and implementation, and even lead the future of clinical practice guidelines has received more attention from researchers all over the world. This paper analyzed and discussed the current situation and prospects of AI in clinical practice guidelines. We also proposed ideas and suggestions on how to promote the integration of AI and guidelines.

**【Key words】** clinical practice guideline; artificial intelligence; machine learning; big data

**Funding:** The National Key R&D Program of China (2018YFC1705500)

*Med J PUMCH*, 2021,12(1):114-121

在医学领域，人工智能（artificial intelligence, AI）目前不仅被广泛应用于辅助诊断、治疗选择、风险预测等<sup>[1]</sup>，还被引入循证医学（evidence-based medicine, EBM）领域，用于系统评价和临床实践指南（下文简称“指南”）的制订、评价、传播与实施等<sup>[2-3]</sup>。虽然目前该领域很多方面仍处于理论探索阶段<sup>[4]</sup>，但进展迅速，未来可期。本文通过分析 AI 在指南中的应用现状，探讨如何通过 AI 缩短指南的制订周期，提高其制订效率，创新其传播与实施模式，并对促进 AI 与指南领域的有效结合提出思考与建议。

1 人工智能在临床实践指南领域的发展

美国食品药品监督管理局迄今已批准 29 款 AI 相关医疗产品<sup>[5]</sup>。尽管 AI 在医学领域的应用取得了一系列较为重要的成果，但其在指南领域的应用仍处于起步阶段。目前国际上 AI 与指南结合的领域主要集中在模型、算法及工具的开发方面，其相关方法已在指南的制订、评价、传播与实施等某些环节小范围内试用，常用的算法和技术包括朴素贝叶斯法和决策树等。表 1 列举了近年来 AI 在指南领域应用的重要事件及进展情况。

我国相关研究人员已在该领域开展积极探索。

2013 年，李博等<sup>[6]</sup>基于语义关系，将高血压指南文本转变为临床实践指南知识库，以提高指南的转化和利用；2019 年，罗旭飞等<sup>[4]</sup>将 AI 在指南领域的应用现状和发展进行了概况评价，结果显示来自美国、澳大利亚、英国和加拿大的研究居多，但大部分研究仍停留于理论探索和算法的验证阶段；2020 年，蒲晓蓉等<sup>[7]</sup>提出了一种基于机器学习的计算机辅助智能分析方法，自动分析不同版本《新型冠状病毒肺炎诊疗方案》的异同。此外，北京大学第三医院开发的 EBM AI-Reviewer（<http://www.ebm-easyreviewer.com/login>），致力于协助系统评价人员快速筛选文献。

2 人工智能与临床实践指南的结合

2.1 人工智能在指南制订中的作用

AI 应用于指南制订领域，旨在缩短指南的制订时间、提高制订质量、优化和加速整个制订流程，现就其中较为重要的 5 个方面展开阐述。

2.1.1 筛选临床问题，确定指南领域

关于研究选题的确定，传统方法主要依靠研究人员的经验。随着 AI 的发展，Chen 等<sup>[8]</sup>采用结构主题模型（structural topic modeling, STM）与文献计量分析相结合的方法，从近十年 AI 辅助人脑研究的 6317

表 1 人工智能在临床实践指南领域应用的重要事件

| 年份     | 重要事件   | 应用环节       |
|--------|--|------------|
| 1996 年 | Ertle 等 <sup>[9]</sup> 应用自动化计数算法，基于指南快速从 EMR 中筛选和提取相关数据，预测和评估哮喘严重程度  | 制订环节       |
| 1998 年 | 哈佛大学、哥伦比亚大学及斯坦福大学的 InterMed 合作项目研究者开发了 GLIF，将指南转化为机器可读格式的模型，实现机构及临床决策支持系统之间的指南共享。GLIF 可在不同医疗机构和系统平台上应用，并与 EMR 系统进行整合。目前，GLIF3 已建立了基于抑郁、流感等编码化指南本体和有效性工具的实例 <sup>[10]</sup> | 评价、传播与实施环节 |
| 1998 年 | Dubey 等 <sup>[11]</sup> 的研究表明 XML 在指南实施方面的使用可能比 HTML 更具优势  | 实施环节       |
| 1999 年 | Mani 等 <sup>[12]</sup> 提出了通过“两阶段”机器学习模型进行数据挖掘、制订指南，并应用于痴呆分期领域中   | 制订环节       |
| 2000 年 | Shiffman 等 <sup>[13]</sup> 提出了 GEM，该模型是一种采用 XML 的指南文档知识模型，可存储和组织临床实践中的异构信息。GEM 定义了 110 个元素，用于表示指南的开发者、目的、目标人群等内容   | 评价、传播与实施环节 |
| 2001 年 | Metfessel 等 <sup>[14]</sup> 开发了提高指南依从性的自动化工具，旨在通过大规模的自动化方法全面追踪指南的规范性，从而推动科学研究向临床实践转化   | 传播与实施环节    |
| 2004 年 | Shahar 等 <sup>[15]</sup> 提出了基于网页、模块化的电子指南库，旨在优化指南的呈现方式和分类汇总形式  | 传播与实施环节    |
| 2007 年 | Kaiser 等 <sup>[16]</sup> 提出了使用自动化提取及格式化指南的工具，并将此方法用于耳鼻咽喉领域的指南  | 评价、传播与实施环节 |
| 2008 年 | Terenziani 等 <sup>[17]</sup> 总结了指南实施及呈现的 AI 方法，并对膀胱癌、反流性食管炎及心力衰竭等疾病进行测试  | 传播与实施环节    |
| 2013 年 | Oliveira 等 <sup>[2]</sup> 从 AI 观点出发，介绍了相关工具，总结了指南的制订和实施  | 制订、传播与实施环节 |
| 2015 年 | Bui 等 <sup>[18]</sup> 开发了帮助检索指南制订证据的机器学习系统   | 制订环节       |
| 2015 年 | 国际系统评价自动化小组成立，并提出了系统评价自动化的维也纳原则 <sup>[19]</sup>  | 制订环节       |
| 2019 年 | Bowles 等 <sup>[20]</sup> 开发了使用自动化方法探测、解决指南中共病药物治疗冲突的框架   | 评价、传播与实施环节 |

EMR：电子病历；GLIF：指南交换格式；XML：可扩展标记语言；HTML：超文本标记语言；GEM：指南元素模型

篇文献中自动识别突出的研究主题，通过对主题趋势、相关性和集群的分析，以及对其在有影响力的国家/地区和研究机构中多样化分布的描述，揭示了有研究前景的主题方向。

关于临床问题的遴选，传统方法一般需耗时数月，对临床一线医务人员进行多轮调查后方可完成，但 AI 技术可对数据库或网络上已有的相关临床问题进行快速分析，自动筛选出高频问题清单，作为指南拟解决问题的重要参考。AI 可对已发表的指南、系统评价及临床研究进行快速、全面的文献扫描和分析，发现指南与研究领域的关联分布，以及指南与临床问题的涵盖关系。

2.1.2 寻找合适专家，管理利益冲突

关于指南专家的遴选，传统方法多采用“方便抽样”，而 AI 有助于查找到最适合参与指南制订的成员，特别是在专业、地域、性别等方面更具代表性。同时，可借助算法分析出每位专家潜在的利益冲突。如 AI 通过对不同渠道信息和数据的关联分析，可发现某专家多次出席某企业产品的推广活动或其直系亲属与医药企业关系密切，指南利益冲突管理委员会可以此作为其存在潜在利益冲突的证据。Graham 等<sup>[21]</sup>

使用机器学习法对 159 种期刊发表的 128 781 篇文章的利益冲突进行分析，发现有商业利益关系的期刊更有可能发表存在利益冲突的文章。

2.1.3 进行研究合成与评价

使用自动化软件可极大缩短系统评价制作过程中文献检索、筛选及质量评价等过程的周期<sup>[22-23]</sup>（表 2）；国际系统评价自动化小组（International Collaboration for the Automation of Systematic Reviews, ICASR）提出的系统评价自动化维也纳原则<sup>[19]</sup>（表 3），为 AI 加速系统评价制作指明了方向。此外，为方便读者选用合适的自动化工具，Marshall 等<sup>[24]</sup>于 2015 年建立了系统评价制作工具箱（systematic review toolbox），用户可登录网站进行文献搜索和选取。AI 技术的应用使系统评价的制作周期缩短至几天甚至几个小时。Clark 等<sup>[22]</sup>使用自动化技术，仅用 2 周时间完成了一篇评估患者增加液体摄入量对泌尿系统感染（urinary tract infection, UTI）的复发、抗菌素的使用及 UTI 症状影响因素的随机对照试验的系统评价。

2.1.4 形成当前最优推荐意见

传统方法制订的指南推荐意见，主要依靠专家对干预措施的安全性、有效性、成本、可实施性及患者

表 2 系统评价中的自动化工具汇总<sup>[22-23]</sup>

| 序号 | 系统评价的制作步骤 | 工具  | 描述或评论  |
|----|-----------|---|--|
| 1  | 制订检索策略    | (1) SRA-Word Frequency Analyser<br>( <a href="https://sr-accelerator.com/#/help/wordfreq">https://sr-accelerator.com/#/help/wordfreq</a> )<br>(2) The Search Refiner<br>( <a href="https://ielab-sysrev2.uqcloud.net/account/login">https://ielab-sysrev2.uqcloud.net/account/login</a> )   | (1) 系统检索时，一般使用高频词汇，SRA-Word Frequency Analyser 通过统计某个词或词组在文章中出现的次数，确定检索策略中应该使用的检索词，如对痛风领域的文献进行分析，发现高频词为“gout”和“hyperuricemia”，检索痛风相关文献时，则需包含上述 2 个词<br>(2) The Search Refiner 通过检查字符串中每个检索词，查找相关研究的数量及不相关研究的数量，快速确定需要删除的检索词，从而加速检索策略的制订                  |
| 2  | 实施检索      | (1) SRA-Polyglot Search Translator<br>( <a href="https://sr-accelerator.com/#/polyglot">https://sr-accelerator.com/#/polyglot</a> )<br>(2) RobotSearch<br>( <a href="https://robotsearch.vortext.systems">https://robotsearch.vortext.systems</a> )<br>(3) Cochrane Register of Studies<br>( <a href="https://community.cochrane.org/help/tools-and-software/crs-cochrane-register-studies">https://community.cochrane.org/help/tools-and-software/crs-cochrane-register-studies</a> )<br>(4) RCT tagger<br>( <a href="http://arrowsmith.psych.uic.edu/cgi-bin/arrowsmith_uic/RCT_Tagger.cgi">http://arrowsmith.psych.uic.edu/cgi-bin/arrowsmith_uic/RCT_Tagger.cgi</a> )<br>(5) Thalia ( <a href="http://nactem-copious.man.ac.uk/Thalia/">http://nactem-copious.man.ac.uk/Thalia/</a> )   | (1) 通过将 PubMed 或 Ovid Medline 搜索转换为其他不同数据库的检索语句以加速检索<br>①经过验证的机器学习过滤器可用于识别 RCT，并可以全自动使用<br>②传统的特定主题关键词搜索策略仍需采用<br>③目前无广泛使用的非 RCT 设计工具<br>(2) 允许在 PubMed 中搜索相关概念（如化学品、疾病、药物、基因、代谢物、蛋白质、物种和解剖实体）   |
| 3  | 去除重复      | SRA-De-duplicator<br>( <a href="https://sr-accelerator.com/#/help/dedupe">https://sr-accelerator.com/#/help/dedupe</a> )  | 从一组上传的记录中识别、删除相同的研究，自动完成大部分重复数据的删除过程；因数据库之间标引存在差异，仍有重复记录存在，此时需手动删除   |
| 4  | 筛选摘要      | (1) RobotSearch ( <a href="https://robotsearch.vortext.systems">https://robotsearch.vortext.systems</a> )<br>(2) SRA Helper ( <a href="https://sr-accelerator.com/#/sra-helper">https://sr-accelerator.com/#/sra-helper</a> )<br>(3) Endnote ( <a href="https://endnote.com/">https://endnote.com/</a> )<br>(4) Abstrackr ( <a href="http://abstrackr.cebm.brown.edu/account/login">http://abstrackr.cebm.brown.edu/account/login</a> )<br>(5) EPPI reviewer ( <a href="https://eppi.ioe.ac.uk/cms/er4">https://eppi.ioe.ac.uk/cms/er4</a> )<br>(6) RobotAnalyst ( <a href="http://www.nactem.ac.uk/robotanalyst/">http://www.nactem.ac.uk/robotanalyst/</a> )<br>(7) SWIFT-Review ( <a href="https://www.sciome.com/swift-review/">https://www.sciome.com/swift-review/</a> )<br>(8) Colandr ( <a href="https://www.colandrapp.com/signin">https://www.colandrapp.com/signin</a> )<br>(9) Rayyan ( <a href="https://rayyan.qcri.org">https://rayyan.qcri.org</a> ) | (1) 从一组检索结果中识别出明显的非 RCT 研究，自动移除这些研究，留下潜在的 RCT 库进行筛选<br>(2) 使用快捷键加速筛选和全文获取，快捷键还可用于搜索预先指定的位置列表，以尝试查找全文<br>(3) 多任务管理，协助管理参考文献，有助于存储搜索结果，查找全文，在筛选过程中进行分组排序，并将参考文献插入文稿中<br>①筛选系统按相关性自动对搜索进行排序<br>②RobotAnalyst 和 SWIFT-Review 还可进行主题建模，与类似主题相关的摘要会自动分组，便于用户检索 |
| 5  | 获取全文      | (1) SRA Helper ( <a href="https://sr-accelerator.com/#/sra-helper">https://sr-accelerator.com/#/sra-helper</a> )<br>(2) SARA ( <a href="https://sr-accelerator.com/#/">https://sr-accelerator.com/#/</a> )  | (1) SRA helper 中快捷键可加速筛选和全文获取，快捷键还可用于搜索预先指定的位置列表，以尝试查找全文<br>(2) 采用 SARA 可自动向图书馆索取全文，只需申请一次即可获取所需全文；通常这些申请需要一次处理并发送；工具可在 SRA 内使用  |
| 6  | 提取数据      | (1) ExaCT<br>( <a href="https://exact.cluster.gctools.nrc.ca/ExactDemo/">https://exact.cluster.gctools.nrc.ca/ExactDemo/</a> )<br>(2) RobotReviewer ( <a href="https://robotreviewer.vortext.systems">https://robotreviewer.vortext.systems</a> )<br>(3) NaCTeM ( <a href="http://www.nactem.ac.uk/software.php">http://www.nactem.ac.uk/software.php</a> )   | 这些软件从自由文本中自动提取数据元素，如样本量大小、PICO 等   |
| 7  | 评估偏倚风险    | RobotReviewer ( <a href="https://robotreviewer.vortext.systems">https://robotreviewer.vortext.systems</a> )   | 通过突出原始 PDF 中的支持性短语，加速评估 7 个“偏倚风险”领域中的 4 个领域；同时结合人工评价对软件评估结果进行验证和检查   |
| 8  | 撰写结果      | RevMan Replicant ( <a href="https://sr-accelerator.com/#/replicant">https://sr-accelerator.com/#/replicant</a> )  | 通过 RevMan 软件中的森林图快速拟定结果初稿，加快结果的撰写  |

SRA：系统评价自动化；RCT：随机对照试验；PICO：患者、干预、对照及结局指标；PDF：便携式文档格式



表 3 系统评价自动化维也纳原则<sup>[19]</sup>

| 条目  | 解释  |
|---|---|
| 原则一：系统评价的制作涉及许多工作，每项工作存在不同问题，但都必须改进                 | 目前证据检索、质量评价、证据综合及发表 4 个步骤可使用自动化工具以提高效率，但需识别出每个步骤可能面临的关键问题                               |
| 原则二：自动化可协助完成系统评价的所有步骤，从确定选题到发现研究空白，从计划书的撰写到全文的撰写及传播 | 标准的系统评价制作一般需 6 个月~2 年，使用自动化方法则可缩短一半时间，目前需关注的是开发能够应用于系统评价整个制作过程的自动化工具                    |
| 原则三：每个步骤都可以且应不断改进，以提高应用自动化方法和工具时的效率和准确性             | 目前大多数自动化工具仍处于开发的早期阶段，因此，应做好这些工具的用户测试和反馈意见收集工作，从而不断改进自动化工具                               |
| 原则四：自动化可以且应高标准地促进系统评价的报告、制作和更新                      | 使用自动化工具和技术的目的是更及时、更省力地制作高质量的系统评价，因此开发和测试新工具时，必须保持较高的标准                                  |
| 原则五：开发商还应提供友好的界面和组件，如将任务进行细分或合并，允许不同用户使用不同的界面       | 不同类型系统评价的制作者需求不同，因此开发自动化工具的过程中，需考虑其用户界面的友好性   |
| 原则六：具有不同专长的工作组在系统评价制作的不同部分发挥作用，需加强小组间的合作            | 系统评价的制作涉及多学科专家的参与，应将资源进行整合，避免不必要的重复工作   |
| 原则七：每一项自动化技术都应该共享，最好免费提供代码、评估数据和语料                  | 共享代码或技术可减少重复研究工作，开发商应向公众提供这些资源，并开放自动化技术的计算代码，以便相关人员可以重复使用并进一步开展工作                       |
| 原则八：所有自动化技术和工具都应使用可靠、可复制的方法进行评估，并报告评估结果和数据          | 自动化过程的所有任务都应由第三方独立评估，评价方法和体例应共享，不仅需评估技术规范，还需评估用户界面等要素，以及其在系统评价制作流程整体的契合度，可防止自我评价可能带来的偏倚 |

意愿等进行综合分析后得出。然而，在指南制订过程中是否对这些因素进行了全面考虑，以及各因素之间相互作用对指南造成影响的系统考量，是传统方法的主要局限性。AI 可以在更广阔的视野下对相关数据进行挖掘和分析，为制订者提供更全面的有效信息，协助制订者更好地平衡干预措施的利弊，形成对患者最优的推荐意见。

2.1.5 改编和更新指南

传统依靠人工更新指南的方法，已满足不了对最新文献的跟进与整合。借助 AI 可推断出最佳证据的出现时间，协助完成指南的及时更新，并确定指南中推荐意见最适合的实施环境。特别是随着动态指南（living guideline）的出现，指南更新的频率和要求越来越高<sup>[25]</sup>。如关于新型冠状病毒肺炎的文献已超过 20 万篇，占 2020 年所有文献的 4%。新型冠状病毒肺炎指南的更新频率已从每月缩短至每周甚至实时<sup>[26]</sup>。AI 可对新文献进行更快速和更高效地监测和遴选；另一方面，还可帮助缺乏指南制订能力的国家改编现有的指南<sup>[27]</sup>。与指南的制订过程不同，指南的改编需评估已有的指南推荐意见，以及与本地区人群和临床环境的匹配性，通过 AI 提供更为全面的当地数据，可协助指南改编人员，作出更合适的推荐意见。

2.2 人工智能在指南传播与实施中的作用

当前指南传播与实施最常用的方法为媒体、

期刊或学术会议宣传等，通过多种途径让使用者知晓指南，进而应用于临床实践。研究显示，这种模式花费时间较长，从指南发布到被指南使用者知晓和应用，需数月甚至数年不等<sup>[28]</sup>。应用 AI 技术，则可将该过程明显缩短。AI 技术将指南中的推荐意见整合至电子病历（electronic medical record, EMR）系统中，通过 OpenMRS 及 DHIS2 等开源 EMR 平台，进一步加速了指南信息的传播，惠及更多医疗资源匮乏的机构或地区<sup>[29-30]</sup>。此外，研究显示，美国约 25% 的医疗费用被浪费掉，每年约合 7600 亿~9350 亿美元<sup>[31]</sup>；若采用 AI 技术，则可以促进指南或医疗信息在患者中的传播，通过去伪存真，为患者提供有价值的信息，从而节约医疗资源。

2.3 人工智能在指南评价中的作用

目前针对单个临床研究的风险偏倚 AI 评估工具已经产生<sup>[32]</sup>。而指南的评价工具，有针对方法学质量的指南研究与评价工具（Appraisal of Guidelines for Research and Evaluation II，AGREE II），包含 6 个领域的 23 个条目<sup>[33]</sup>；以及针对报告质量的国际实践指南报告标准（Reporting Items for Practice Guidelines in Healthcare, RIGHT），包含 7 个领域的 22 个条目<sup>[31]</sup>。然而，研究显示使用这些工具时存在一定挑战，即不同的使用者对标准的把握不一致，人力和时间成本亦不可控<sup>[34-35]</sup>。如使用 AGREE II 评价一部指南，

常规情况下需要 3~4 名经过培训的研究人员，每人至少耗时 1 h；如评价糖尿病领域百余部指南的质量，则需 500 h 以上。但通过开发、训练相应的 AI 算法，不仅可以降低评价的主观偏倚，还可较大程度上缩短评价所需的时间。此外，对指南的评价还包括对推荐意见的对比分析，而快速提取其中的推荐意见是关键步骤。Hussain 等<sup>[36]</sup>开发的 AI 算法，用于提取高血压指南中的推荐意见，其准确度可达 85.54%。

3 促进人工智能与临床实践指南结合的思考与建议

就目前科技进步的速度和学科交叉的深度而言，AI 技术与指南的结合已是箭在弦上。指南的制订已从 1.0 时代迈向 2.0 时代<sup>[37]</sup>，而 AI 很有可能引领指南发展的 3.0 时代。为提前做好应对并促进其科学发展，笔者提出以下思考与建议。

3.1 政策制定者与管理者层面

AI 在指南领域的发展和应用，除技术支持外，还有赖于政策支撑。2021 年，国家自然科学基金委新增设了交叉科学部，为 AI 与指南结合相关研究工作的开展提供了获取更多支持的可能性。近年来，“国家重点研发计划”“科技创新 2030-重大项目”等也加大了对 AI、大数据在医学领域应用的支持力度。此外，多所高校先后设立了人工智能学院，开设 AI 与医学交叉相关课程，招收本科生和研究生，并在基础理论和关键技术等方面开展相关研究和培训，是未来在该领域把握机遇和掌握话语权的重要途径。

3.2 指南制订者和实施者层面

在指南制订者和实施者层面主要提出以下建议：（1）制订指南或开展指南研究时，应纳入 AI 相关学科人员。（2）使用已被证实可靠的 AI 算法、模型和工具时，应提供相应的指导手册或操作规范。（3）撰写基于 AI 的指南手册或专著。英国国家健康与临床优化研究所（National Institute for health and Clinical Excellence, NICE）在其指南制订手册中纳入应用机器学习搜索和筛选文献的内容<sup>[38]</sup>，其他指南制订手册则无相关内容的描述。（4）借助 AI 加速指南的传播与实施，促进医患沟通，改善患者关系。正如斯坦福大学 AI 实验室主任 Christopher Manning 教授所言：“AI 可以把指南带入每一次的患者咨询中，让每一次咨询都安全、稳

妥地支持医患双方沟通。”<sup>[39]</sup>

3.3 研究者层面

在研究者层面主要提出以下建议：（1）研发基于 AI 的系统评价工具。指南制订的限速环节主要是系统评价的制作，尽管当前已有 AI 协助系统评价制作的案例，但仍需进一步深入研究，推动 AI 大范围应用于系统评价的制作是当前指南制订最重要和最需优先开展的研究。（2）研发 AI 相关的报告规范。指南制订过程中应用 AI 的关键步骤和信息应予以充分、透明地报告，故应研发 AI 应用于指南的报告规范“RIGHT for AI”，现已成为 RIGHT 工作组扩展版内容之一。目前，AI 声明“CONSORT（Consolidated Standards of Reporting Trail）for AI”已完成研发，其成果已在学术期刊上发表<sup>[40]</sup>。随着 AI 相关系统评价的不断增多，系统评价和 Meta 分析优先报告条目“PRISMA（Preferred Reporting Items for Systematic Reviews and Meta-Analyses）for AI”的出现将是必然。在 AI 相关的临床研究、系统评价和指南等领域形成完整的报告规范体系，对于促进 AI 发展、彰显其应用价值具有重要意义。（3）开展 AI 识别与管理指南参与人员的利益冲突、患者偏好及形成推荐意见等相关研究是未来需要探索的重要方向。如利用 AI 收集患者偏好与价值观等信息时，如何考虑公平性（当前关于患者偏好与价值观的研究基本上是基于西方患者人群）；管理指南参与人员的利益冲突时，如何保护其个人隐私等，都是研究者需要关注和解决的问题。AI 为指南提供的信息越全面，制订者做出最优推荐意见的可能性越大。

4 小结

目前，AI 在指南领域的应用仍处于起步阶段，未来前景广阔，其可在指南的制订、实施到自动更新全过程中发挥作用，但也面临诸多问题与挑战，这些问题和挑战将随着人们对 AI 的认识和掌握不断加深而转化为机遇。我们可以期待，未来指南的制订也许只需基于患者个体数据便可产出极具针对性的“基于证据的个体化指南”，届时精准医学的受益对象将不再是少数人的罕见病，而是所有人的大健康。

**作者贡献：**陈耀龙、罗旭飞负责统计数据并撰写文章初稿；史乾灵、吕萌、周奇、王健健、杨楠、高东平、杨书、商洪才、杨克虎负责修改初稿；陈耀龙负责文章审校。

利益冲突：无

## 参 考 文 献

- [1] He JX, Baxter SL, Xu J, et al. The practical implementation of artificial intelligence technologies in medicine [J]. *Nat Med*, 2019, 25: 30-36.
- [2] Oliveira T, Novais P, Neves J. Development and implementation of clinical guidelines: An artificial intelligence perspective [J]. *Artif Intell Rev*, 2014, 42: 999-1027.
- [3] Isern D, Moreno A. Computer-based execution of clinical guidelines: a review [J]. *Int J Med Inform*, 2008, 77: 787-808.
- [4] Luo XF, Xiao YJ, Liu X, et al. How to Combine Artificial Intelligence with Clinical Practice Guideline: A Scoping Review[C/OL].(2019-10-31). [https://journals.lww.com/ijebh/fulltext/2019/12000/abstracts\\_for\\_the\\_gin\\_and\\_jbi\\_conference,\\_adelaide.2.aspx](https://journals.lww.com/ijebh/fulltext/2019/12000/abstracts_for_the_gin_and_jbi_conference,_adelaide.2.aspx).
- [5] Benjamins S, Dhunoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database [J]. *NPJ Digit Med*, 2020, 3: 118.
- [6] 李博, 李科, 曾东, 等. 基于语义关系的高血压临床指南知识库构建 [J]. *中国数字医学*, 2013: 64-67.
- [6] Li B, Li K, Zeng D, et al. Construction of Hypertension Clinical Guideline Knowledge Base Based on Semantic [J]. *Zhongguo Shu Zi Yi Xue*, 2013: 64-67.
- [7] 蒲晓蓉, 陈柯成, 刘军池, 等. COVID-19 中国诊疗方案的机器学习智能分析方法 [J]. *生物医学工程学杂志*, 2020, 37: 365-372.
- [7] Pu XR, Chen KC, Liu JC, et al. Machine learning-based method for interpreting the guidelines of the diagnosis and treatment of COVID-19 [J]. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi*, 2020, 37: 365-372.
- [8] Chen XL, Chen J, Cheng G, et al. Topics and trends in artificial intelligence assisted human brain research [J]. *PLoS One*, 2020, 15: e0231192.
- [9] Ertle AR, Campbell EM, Hersh WR. Automated application of clinical practice guidelines for asthma management [J]. *Proc AMIA Annu Fall Symp*, 1996: 552-556.
- [10] Ohno-Machado L, Gennari JH, Murphy SN, et al. The guideline interchange format: a model for representing guidelines [J]. *J Am Med Inform Assoc*, 1998, 5: 357-372.
- [11] Dubey AK, Chueh H. Using the extensible markup language (XML) in automated clinical practice guidelines [J]. *Proc AMIA Symp*, 1998: 735-739.
- [12] Mani S, Shankle WR, Dick MB, et al. Two-stage machine learning model for guideline development [J]. *Artif Intell Med*, 1999, 16: 51-71.
- [13] Shiffman RN, Karras BT, Agrawal A, et al. GEM: a proposal for a more comprehensive guideline document model using XML [J]. *J Am Med Inform Assoc*, 2000, 7: 488-498.
- [14] Metfessel BA. An automated tool for an analysis of compliance to evidence-based clinical guidelines [J]. *Stud Health Technol Inform*, 2001, 84: 226-230.
- [15] Shahar Y, Young O, Shalom E, et al. The Digital electronic Guideline Library (DeGeL): a hybrid framework for representation and use of clinical guidelines [J]. *Stud Health Technol Inform*, 2004, 101: 147-151.
- [16] Kaiser K, Akkaya C, Miksch S. How can information extraction ease formalizing treatment processes in clinical practice guidelines? A method and its evaluation [J]. *Artif Intell Med*, 2007, 39: 151-163.
- [17] Terenziani P, Montani S, Bottrighi A, et al. Applying artificial intelligence to clinical guidelines: the GLARE approach [J]. *Stud Health Technol Inform*, 2008, 139: 273-282.
- [18] Bui DD, Jonnalagadda S, Del Fiore G. Automatically finding relevant citations for clinical guideline development [J]. *J Biomed Inform*, 2015, 57: 436-445.
- [19] Beller E, Clark J, Tsafnat G, et al. Making progress with the automation of systematic reviews: principles of the International Collaboration for the Automation of Systematic Reviews (ICASR) [J]. *Syst Rev*, 2018, 7: 77.
- [20] Bowles J, Caminati MB, Cha S, et al. A framework for automated conflict detection and resolution in medical guidelines [J]. *Sci Comput Program*, 2019, 182: 42-63.
- [21] Graham SS, Majdik ZP, Clark D, et al. Relationships among commercial practices and author conflicts of interest in biomedical publishing [J]. *PLoS One*, 2020, 15: e0236166.
- [22] Clark J, Glasziou P, Del Mar C, et al. A full systematic review was completed in 2 weeks using automation tools: a case study [J]. *J Clin Epidemiol*, 2020, 121: 81-90.
- [23] Marshall IJ, Wallace BC. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis [J]. *Syst Rev*, 2019, 8: 163.
- [24] Marshall C, Brereton P. Systematic review toolbox: a catalogue of tools to support systematic reviews [C/OL].(2015-05-08). [https://www.researchgate.net/publication/275967185\\_Systematic\\_Review\\_Toolbox\\_A\\_Catalogue\\_of\\_Tools\\_to\\_Support\\_Systematic\\_Reviews](https://www.researchgate.net/publication/275967185_Systematic_Review_Toolbox_A_Catalogue_of_Tools_to_Support_Systematic_Reviews).
- [25] Akl EA, Meerpohl JJ, Elliott J, et al. Living Systematic Re-

- view Network. Living systematic reviews; 4. Living guideline recommendations [J]. J Clin Epidemiol, 2017, 91: 47-53.
- [26] Else H. How a torrent of COVID science changed research publishing-in seven charts [J]. Nature, 2020, 588: 553.
- [27] Groot P, Hommersom A, Lucas P. Adaptation of clinical practice guidelines [J]. Stud Health Technol Inform, 2008, 139: 121-139.
- [28] Gagliardi AR, Marshall C, Huckson S, et al. Developing a checklist for guideline implementation planning: review and synthesis of guideline development and implementation advice [J]. Implement Sci, 2015, 10: 19.
- [29] Landis Lewis Z, Mello-Thoms C, Gadabu OJ, et al. The feasibility of automating audit and feedback for ART guideline adherence in Malawi [J]. J Am Med Inform Assoc, 2011, 18: 868-874.
- [30] Wahl B, Cossy-Gantner A, Germann S, et al. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? [J]. BMJ Glob Health, 2018, 3: e000798.
- [31] Shrank WH, Rogstad TL, Parekh N. Waste in the US Health Care System: Estimated Costs and Potential for Savings [J]. JAMA, 2019, 322: 1501-1509.
- [32] Marshall IJ, Kuiper J, Wallace BC. RobotReviewer: evaluation of a system for automatically assessing bias in clinical trials [J]. J Am Med Inform Assoc, 2016, 23: 193-201.
- [33] Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care [J]. CMAJ, 2010, 182: E839-E842.
- [34] Chen YL, Yang KH, Marušić A, et al. A reporting tool for practice guidelines in health care: the RIGHT statement [J]. Ann Intern Med, 2017, 166: 128-132.
- [35] Tokalić R, Viđak M, Buljan I, et al. Reporting of Clinical Practice Guidelines: Practical Testing of AGREE and RIGHT Checklists [J]. J Gen Intern Med, 2020, 35: 2167-2172.
- [36] Hussain M, Hussain J, Sadiq M, et al. Recommendation statements identification in clinical practice guidelines using heuristic patterns [C/OL]. (2018-07-27). <https://www.computer.org/csdl/proceedings-article/snspd/2018/08441036/13bd1sv5NyA>.
- [37] Schünemann HJ, Wiercioch W, Etzeandía I, et al. Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise [J]. CMAJ, 2014, 186: E123-E142.
- [38] National Institute for Health and Care Excellence. Developing NICE guidelines: the manual [M/OL]. (2020-10-15). <https://www.nice.org.uk/process/pmg20/chapter/introduction>.
- [39] Manning CL. Artificial intelligence could bring relevant guidelines into every consultation [J]. BMJ, 2019, 366: l4788.
- [40] Liu XX, Rivera SC, Moher D, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension [J]. BMJ, 2020, 370: m3164.

(收稿: 2021-01-05 录用: 2021-01-17)

(本文编辑: 李玉乐)