

中医临床决策支持系统研发的现状与思考

白宇¹ 陈理² 易岚² 胡荣魁² 王石² 范群丽² 李文友³ 赵静³
韩捷² 罗倩倩² 陈晓虎²

(1.南京中医药大学中医学院·中西医结合学院,江苏南京210023;2.南京中医药大学附属医院,江苏南京210029;
3.南京大经中医药信息技术有限公司,江苏南京211800)

摘要 结合了人工智能和大数据的临床决策支持系统(CDSS)可以在医疗的各个环节和方面为临床医生提供决策支持。近年来,随着市场对CDSS的需求越来越大,研发临床适用性高的CDSS来提高医疗质量已成为热门趋势。中医CDSS的研发起步较晚且难度远高于西医,虽然目前已取得了一定的成果,但是在模拟中医思维、数据录入标准、获取临床数据等方面存在比较大的缺陷,难以提供良好的用户体验。在未来的研发过程中,应当进一步加强中医术语的规范化研究,以古籍和名老中医病案为主要数据来源,从单病种开始逐渐突破,最终形成面向中医各个学科的CDSS。

关键词 临床决策支持系统;人工智能;机器学习;中医思维

基金项目 江苏省科技计划项目(重点项目-临床前沿技术,SBE2019750375)

临床决策支持系统(clinical decision support system, CDSS)是通过数据、模型和算法等智能化的手段,以人机交互方式来辅助决策者的计算机应用系统^[1]。它可以模拟医学专家诊断、治疗疾病的思维和过程,在疾病诊断、治疗、护理、手术、合理用药等方面为临床医生提供决策支持。近年来,运用CDSS提高医疗质量,辅助临床教学,助力科学研究已成为主流趋势。

随着人口老龄化的加剧,慢病、疑难病的高速增长,使得我国对医疗人工智能需求巨大。但我国CDSS的研发明显滞后于国际先进水平,尤其在中医领域更是如此。针对这一现状,国务院办公厅于2018年发布《关于促进“互联网+医疗健康”发展的意见》,提出“支持中医辨证论治智能辅助系统应用,提升基层中医诊疗服务能力”。但由于种种问题的存在,目前仍未设计出具备较为理想的临床实用性的中医CDSS。本文就目前中医CDSS研发面临的问题和解决方法进行相关阐述与讨论。

1 中医CDSS的研究现状

1.1 中医临床思维难以用计算机进行模拟 中医学的思维源于中国古代哲学,其核心的思维模式是象思维,即以物象或意象为解析工具,把客观事物自然整体显现于外的现象作为依据,运用比喻、象征、类

推等方法表达对象世界的抽象意义,以此把握对世界的普遍联系乃至本原之象^[2]。象思维具体表现为“整体观念”和“辨证论治”^[3]。“整体观念”是中医思维的核心,贯穿于中医的各个方面,而“辨证论治”则是“整体观念”在临床诊治过程的具体应用,是中医实现动态、个性化诊治的基础。在两者的指导下,不同医家结合长期的临床实践,对疾病发生发展的规律出现了不同的理解,又逐步建立了“八纲”“脏腑”“病因”“六经”“卫气营血”和“三焦”等不同的辨证方法。这些辨证方法本身就各有其自身的特点,再加上个人体质、环境气候、饮食习惯等多种因素影响,又会出现难以计数的个体差异,使得中医的辨证思维模式非常难以复刻。

中医CDSS的技术架构需要通过数据层、知识层(知识图谱层)、AI层、应用层来逐层实现,其中知识层和AI层是中医CDSS的核心环节。知识层依托数据层通过对数据的进一步抽象和整理,形成知识图谱,让机器理解中医的知识和思维。AI层则是通过自然语言处理和机器学习等算法,建设不同的模型以满足个性化的智能需求,是对数据和知识的应用能力的体现。就目前而言,CDSS涉及的人工智能算法包括K近邻算法(k-nearest neighbor, KNN)、回归算法、决策树、随机森林、贝叶斯决策

论^[4]等。在算法之上又有机器学习和神经网络技术综合各类学科和算法,实现计算机对人类的学习行为和大脑思维模式的模拟,从而获取新知识和技能,并重新优化已有的知识结构使之不断改善自身^[1]。虽然这些方法对中医由知识和经验转化成数据提供了多层面的技术支持^[5],但即使如此,在面对中医思维模式时,这些技术仍然显得不够成熟。除此以外,中医思维模式常常被认为是“黑箱”理论的一种体现^[6-7],但是在CDSS的设计中“黑箱”理论的应用是存在争议的^[8]。在CDSS的设计理念中,CDSS做出的每个决策都必须保证能给出透明的推理演算过程,以使用户能够理解所有CDSS提供的建议的原理。这也在无形之中给中医人工智能之路增加了不小的困难。

1.2 数据录入缺乏统一的标准 和其他领域一样,中医临床数据同样可分为结构化数据、非结构化数据以及半结构化数据。结构化数据是指具有一定逻辑结构和物理结构关系,可以存储在数据库中的数据,也被称为定量数据,如年龄、性别及生化检测指标等。非结构化数据是相对于结构化数据而言的,这类数据不存储在数据库中,而是以各种类型的文本形式存放,又称为文本数据,如影像学中的各类图像。半结构化数据是介于结构化数据和非结构化数据之间的数据,最主要的就是病历资料。

中医学涉及的大量病案、古籍及四诊信息大都属于半结构化数据或非结构化数据^[9]。半结构化数据或非结构化数据不能被电子病历系统(electronic medical record, EMR)直接利用,必须经过一定的转换使其能够被计算机识别。尽管现在自然语言处理(natural language processing, NLP)可以协助完成知识提炼和知识提取,最终形成知识本体或者知识网络^[10-11],但前提是必须建立一套完善的中医术语库。这套中医病案症状术语库必须满足两个条件:一是能准确且真实地表达中医思想,否则会诱导人工智能产生偏差;二是必须形成一定的应用规模,否则将无法实现数据整合,形成一片片“信息孤岛”^[12]。中医医案的记录主观因素过多、随意性强且结构复杂,各类症状的内涵和外延模糊^[13],各位医家对同样的症状会产生完全不同的理解和描述,因此尚未建立被广泛接受并使用的中医术语库。

除了症状以外,如何对中医治疗进行疗效评价也是一个亟待解决的问题。因为如果没有明确的中医疗效评价标准,机器学习就无法对既有的算法和模型进行反馈,从而无法自行进一步学习优化。目

前中医疗效评价方法主要有两种:一是利用证候积分量表,二是通过观察治疗前与治疗后病机的变化判断疗效^[14]。无论哪种都依靠医生的经验与主观感受,依赖患者症状与体征的变化进行疗效评价,会带来重复性差、主观性强并且观察时间有限等问题^[15]。但如果中医疗效的评价方法和指标只是单纯照搬生物医学的模式,又很难反映出中医的自身规律与特点,反而会对中医药的持续发展产生不利的影响。

1.3 缺乏高质量临床数据 大数据需要人工智能,人工智能也需要大数据。战胜李世石和柯洁的阿尔法狗(AlphaGo)学习了超过三千万盘棋局,而作为一个中医CDSS系统保守估计至少需要上千份有质量的病历资料来进行机器学习。虽然目前已有微医集团的“悬壶台”、朱文锋教授的“WF文锋-Ⅲ中医(辅助)诊疗系统”等多种中医CDSS问世,但由于缺乏一套完备可靠的中医临床数据来源,无法真实模拟中医诊疗疾病思维的过程,大大限制其辅助诊疗的正确性和临床实用性,无法提供高质量且合理的管理建议,无法达到理想的用户满意度和实用性。当前大部分医疗数据,特别是中医数据都是准确度低且缺乏完整性的非结构化无效数据,因为在临床病例采集中,即使对于同一患者,不同采集者所采集的数据都难以做到一致。而中医更是具有非常个性化的特点,致使中医数据的产生具备了高度的不可预测性,这也严重影响中医大数据的利用。

2 方法与对策

2.1 加速中医药规范化进程,符合信息化需求 目前我国已投入大量人力物力进行中医规范化术语研究,且已经取得了一些成果,建立了一些具有一定规模的术语库,如中医药学语言系统(TCMLS)、中医临床术语系统、中医古籍语言系统等许多大型的术语系统^[16]。但是,这些标准仍未完全符合国际标准化组织技术规范,仅可实现简单的检索功能,难以进行深度数据挖掘与分析^[17]。2019年中国中医药信息学会发布了57项团体标准(T/CIATCM 001~057—2019),是我国中医药信息行业的首批团体标准。这57项标准涵盖中医药信息的基础标准、数据标准、技术标准和管理标准等101项中医药信息标准研究项目。其中《中医临床基本症状信息分类与代码》中包含了881个中医基本症状信息的分类代码表,基本能够满足中医临床病历的需求。虽然这个团体标准与临床应用之间还有一定的距离,但是相信通过不断的临床应用和反馈,最终必然会形成一套完善的可以应用于大数据研究的中医药规范化术语库。

2.2 利用名老中医病案和中医古籍,提高数据质量 中医药大数据研发缺乏可靠的临床数据,这是短时间内无法解决的难题。相较于大量质量参差不齐的中医临床数据,名老中医的临床病案和中医古籍中记载的经典方药,在质量上更适合直接被人工智能采用。

名老中医的诊疗经验传承是中医药理论创新发展的思想源泉,也是发掘我国中医药宝藏的重要学术支撑。截至目前,国家已开展了六批3600余人次的全国老中医药专家学术经验继承工作、1200余个名老中医传承工作室建设工作,已拥有200多位名医的医案4万份、典型医案6000余份、有效方药300多首等数据资源。这些珍贵的特色病案与有效方药不仅是中医最宝贵的财富,也是可靠的中医临床数据来源。尽管传承成果丰硕,但由于名老中医诊疗经验的传承模式目前仍以师带徒为主,效率低下且周期过长,缺乏现代科技的支撑。因此如何将名老中医临床病案进行数据处理,挖掘其中隐含的诊疗知识并进行总结研究,进而模拟名老中医的诊疗思维,研发基于名老中医诊疗经验的CDSS已成为新的研究热点。这种结合模式不仅能解决中医CDSS研发缺乏临床大数据支持的问题,同时也能解决名老中医传承科技水平低下、缺乏成果转化问题。

与名老中医经验相似,中医古籍中记载的经典方是中医理论和诊疗实践传承的载体,蕴含着巨大的信息量。据2007年上海辞书出版社出版的《中医古籍总目》记载,全国131个图书馆中馆藏中医书目有1.3万种,涉及范围包括临床诊疗经验、中药采集、药性辨别、医案解释和偏方验方汇集等。屠呦呦正是通过古籍文献带来的灵感发现了青蒿素,于2015年10月获得诺贝尔生理学或医学奖,这说明古籍中蕴含着极为有价值的信息。目前对中医古籍的数据挖掘大多比较粗浅,主要是一些用药频次研究或简单的关联分析^[18],但也曾有学者将人工神经网络、决策树等人工智能手段融入到古籍数据挖掘中^[19]。虽然这些研究成果的临床实用性还有待验证,但是足以说明对中医古籍进行数据挖掘,从而为中医CDSS提供基础数据和机器学习的素材是一条潜在的可行之路。

2.3 从单病种开始,逐渐完成融合 以目前的技术和前期研究成果,想要一步到位研发出面向各个专科的中医CDSS几乎是不可能的事。但是我们可以先从单病种开始,研发出针对某一专科的CDSS系统,最终

融合形成全面的面向中医各个学科的CDSS系统。回顾CDSS在国外的发展趋势,我们发现即使在CDSS研发起步较早的美国,目前仍然是以专科CDSS研发为主^[20]。例如用于精神病学心理健康管理决策支持的CEMS (clinical evaluation and monitoring system) 系统,用于传染病诊断决策支持的全球传染病和流行病学网络(global infectious diseases and epidemiology network, GIDEON) 系统^[21],用于妇产科临床决策支持的IPROB系统及用于艾滋病病毒(HIV)感染患者药物治疗决策支持的RetroGram系统^[22,20],等等。虽然也有例如DXplain和伊莎贝尔(Isabel)系统在内的运用范围较广的CDSS出现,但应用场景仅包括案例分析、教学辅助、定点照护及医疗咨询^[23],距离理想中的CDSS系统还相去甚远。

中医CDSS的研发难度远高于西医,因此我们也不必急于研发面向各个专科的中医CDSS,可以先着手研发针对单个中医专科的CDSS,等时机成熟了以后再将其融合,形成完整的中医CDSS。这样可以暂时降低研发的难度,为后期的研发提供宝贵的经验,反而有可能加快CDSS的研发速度,减少走弯路的过程,不失为一种较为有效的探索方式。

3 结语

自1950年逻辑健康处理评价(health evaluation through logical processing, HELP)系统问世以来,CDSS的研发已经走过了70多年的历程。我国自主研发的CDSS也如雨后春笋般涌现,并逐步投入临床应用,腾讯和阿里等互联网公司均涉足其中。但限于数据、算法等制约因素,基于人工智能的中医辅助诊疗系统发展非常滞后,目前依然没有人工智能在中医药领域应用的成功案例。虽然困难重重,但是市场对中医CDSS的需求却一直非常强大。仅从慢性病防治这一应用层面来说,中医CDSS就具有协助广大青年医生完成临床决策,提高广大基层医务工作者慢病防治水平,提升医疗质量,控制医疗费用支出,提高社会资源利用率的作用。

中医的发展需要传承与创新。基于时代高速发展的人工智能技术为中医的传承和创新提供了非常好的途径和路径,使得中医的创新形式更加多样化和标准化,与现代医学的结合也更强。尽管中医CDSS在多种方面存在很多问题与挑战,但凭借着广阔的发展空间和应用前景,相信在未来中医人工智能的帮助下可以解决中医发展过程中面临的更多实际问题,为人类健康作出更多贡献。

参考文献

- [1] GARCIA-VIDAL C, SANJUAN G, PUERTA-ALCALDE P, et al. Artificial intelligence to support clinical decision-making processes[J]. EBioMedicine, 2019, 46 : 27.
 - [2] 邢玉瑞. 中医象思维的概念[J]. 中医杂志, 2014, 55 (10) : 811.
 - [3] 赵文, 林雪娟, 闵莉, 等. 中医思维的内涵与外延[J]. 中华中医药杂志, 2020, 35 (1) : 46.
 - [4] 王慧, 冀晓亮. 鸢尾花数据集剖析人工智能经典算法[J]. 科技与创新, 2021 (18) : 14.
 - [5] 郭仪, 许斌, 胡楠. 人工智能在辅助中医临床领域的研究现状与展望[J]. 中华中医药学刊, 2021, 39 (6) : 76.
 - [6] 关晓光, 王晓鹏, 栾广君. 中医司外揣内与黑箱方法: 自然辩证法概论教学案例之一[J]. 中国中医药现代远程教育, 2009, 7 (12) : 80.
 - [7] 钱丽. 黑箱方法与中医“藏象学说”[J]. 南京中医药大学学报(社会科学版), 2004, 5 (1) : 14.
 - [8] SHORTLIFFE E H, SEPÚLVEDA M J. Clinical decision support in the era of artificial intelligence[J]. JAMA, 2018, 320 (21) : 2199.
 - [9] 邵明义, 刘保延, 谢琪, 等. 中医药临床科研数据的发展现状和趋势探讨[J]. 世界科学技术-中医药现代化, 2015, 17 (8) : 1743.
 - [10] 屈丹丹, 杨涛, 胡孔法. NLP在中医医案症状信息自动化抽取中的应用研究[J]. 软件导刊, 2021, 20 (2) : 44.
 - [11] 柴华, 路海明, 刘清晨. 中医自然语言处理研究方法综述[J]. 医学信息学杂志, 2015, 36 (10) : 58.
 - [12] 王海星, 田雪晴, 游茂, 等. 人工智能在医疗领域应用现状、问题及建议[J]. 卫生软科学, 2018, 32 (5) : 3.
 - [13] 谢蓉, 王燕萍, 彭丹虹, 等. 中医症状规范化研究[J]. 河南中医, 2017, 37 (7) : 1144.
 - [14] 石强, 张倩. 建立分层次中医疗效评价体系的思考[J]. 医学食疗与健康, 2020, 18 (24) : 188.
 - [15] 何泽慧, 张戈, 欧爱华. 关于中医临床疗效评价研究的思考[J]. 中国民族民间医药, 2020, 29 (1) : 115.
 - [16] 于彤, 崔蒙, 李海燕, 等. ISO技术规范“中医药学语言系统语义网络框架”的应用研究[J]. 中国医药导报, 2016, 13 (4) : 89.
 - [17] 朱彦, 朱玲, 王俊慧, 等. 基于信息抽取的历代方剂药物知识发现方法及应用[J]. 中华中医药杂志, 2015, 30 (5) : 1447.
 - [18] 曾大方. 《临证指南医案》方案的统计研究[J]. 山东中医学院学报, 1985 (1) : 13.
 - [19] 周德生. 明清时期津液亏损病案573例辨证用药统计分析[J]. 中医药研究, 1998 (4) : 12.
 - [20] 盛赞, 张越. 美国临床决策支持系统发展与启示[J]. 中国卫生信息管理杂志, 2016, 13 (3) : 257.
 - [21] KEYSTONE J S. GIDEON computer program for diagnosing and teaching geographic medicine[J]. J Travel Med, 1999, 6 (2) : 152.
 - [22] TURAL C, RUIZ L, HOLTZER C, et al. Clinical utility of HIV-1 genotyping and expert advice: the Havana trial[J]. AIDS, 2002, 16 (2) : 209.
 - [23] VARDELL E, MOORE M. Isabel, a clinical decision support system[J]. Med Ref Serv Q, 2011, 30 (2) : 158.
- 第一作者: 白宇 (1988—), 男, 医学博士, 讲师, 主治中医师, 从事中西医结合临床及基础研究。
 通讯作者: 陈晓虎, 医学博士, 教授, 主任中医师, 博士研究生导师。cxhjsszyy@126.com
 修回日期: 2022-10-22
 编辑: 吴 宁

文末参考文献著录规则之页码的标注

根据国家标准GB/T 7714—2015《信息与文献 参考文献著录规则》的规定, 本刊关于参考文献页码的著录做重要修订如下: (1) 专著或期刊中析出文献的页码或引文页码, 应采用阿拉伯数字著录, 引自序言或扉页题词的页码, 可按实际情况著录 (例: 钱学森. 创建系统学[M]. 2版. 太原: 山西科学技术出版社, 2001: 序. 2.); (2) 阅读型参考文献的页码著录文章的起始页, 引文参考文献的页码著录引用信息所在页。阅读型参考文献指著者为撰写或编辑论著而阅读过的信息资源; 引文参考文献指著者为撰写或编辑论著而引用的信息资源。例如: 将谈勇等发表在本刊2015年第1期第1—4页的文章《夏桂成国医大师调治复发性流产经验探赜》作为阅读型参考文献引用时, 页码著录为“1”; 作为引文文献引用“夏老提出心-肾-子宫轴功能失常是流产病机关键”这一观点或原文时, 著录这些引用信息的所在页“3”。更多有关本刊参考文献著录规则见本刊网站 (www.jstem.cn) 首页下载专区。