

传统中药标准的数字化新形态

王淑红^{1,2}, 李君瑶^{1,2}, 郭子瑜^{1,2}, 王子骏³, 何轶^{4,5}, 苏畅^{1,2}, 谷丽菲^{1,2}, 卢光明³, 王冰^{1,2*}, 马双成^{4,5*}

(1. 深圳市药品检验研究院, 广东 深圳 518057; 2. 国家药品监督管理局中药质量研究与评价重点实验室, 广东 深圳 518057; 3. 哈尔滨工业大学(深圳), 广东 深圳 518055; 4. 国家药典委员会, 北京 100061; 5. 药品监管科学全国重点实验室, 北京 100061)

摘要:目的 数字时代下,中药标准数字化转型成为提升中药质量控制,推动产业高质量发展的必然趋势。**方法** 笔者研究团队利用 AI 技术在图像识别分析、数据挖掘处理、模型预测及决策支持等方面的优势,开展传统中药标准的数字化转型。**结果与结论** 尝试构建一种无人工干预的机器可读、可执行、可解析、可决策的中药标准数字化新形态,积极推动中药标准的数字化、网络化、智能化发展,服务标准数字化水平不断提高的国家战略目标。

关键词: 中药数字标准; 中药标准; 数字化转型; 新质生产力; 数字技术

doi:10.11669/cpj.2024.24.013 中图分类号:R282 文献标志码:A 文章编号:1001-2494(2024)24-2396-11

New Digital Form of Traditional Chinese Medicine Standard

WANG Shuhong^{1,2}, LI Junyao^{1,2}, GUO Ziyu^{1,2}, WANG Zijun³, HE Yi^{4,5}, SU Chang^{1,2}, GU Lifei^{1,2}, LU Guangming³, WANG Bing^{1,2*}, MA Shuangcheng^{4,5*} (1. Shenzhen Institute for Drug Control, Shenzhen 518057, China; 2. NMPA Key Laboratory for Quality Research and Evaluation of Traditional Chinese Medicine, Shenzhen 518057, China; 3. Harbin Institute of Technology, Shenzhen 518055, China; 4. Chinese Pharmacopoeia Commission, Beijing 100061, China; 5. State Key Laboratory of Drug Regulatory Science, Beijing 100061, China)

ABSTRACT:OBJECTIVE In the digital era, the digital transformation of traditional Chinese medicine standards has become an inevitable trend to improve the quality control of traditional Chinese medicine and promote the high-quality development of the industry.

METHODS The research team has leveraged AI technology's advantages in image recognition and analysis, data mining and processing, model prediction, and decision support to carry out the digital transformation of traditional Chinese medicine standards.

RESULTS AND CONCLUSION They aim to construct a new form of traditional Chinese medicine digital standards that is machine-readable, executable, interpretable, and decision-capable without human intervention. This effort actively promotes the digitalization, networking, and intelligent development of traditional Chinese medicine standards, contributing to the national strategic goal of continuously improving the level of standard digitization.

KEY WORDS: traditional Chinese medicine digital standard; traditional Chinese medicine standard; digital transformation; new quality productive force; digital technology

中药标准作为中药质量控制方法的重要载体,是药品监管的重要工具。2023年,国家药品监督管理局关于《进一步加强中药科学监管促进中药传承创新发展若干措施》(国药监药注〔2023〕1号)要求“提升中药标准数字化管理水平。建立完善中药国家药品标准、药品注册标准动态数据库,加快推进数字化标准建设,及时更新数据,实现药品标准的发布、查询、分析、研究、维护信息化^[1]。”中药标准迈向数字化转型新阶段。

1 全球标准数字化进展

标准是经济活动和社会发展的技术支撑,是国家基础性制度的重要方面^[2]。近年来,在全球范围内,国际标准化组织和欧洲、美国及亚洲的主要经济体均围绕标准数字化及相关主题开展研究和实践。

1.1 数字化政策现状

1.1.1 国家数字化政策^[2-3] 2021年10月,中共中央、国务院印发了《国家标准化发展纲要》,将“标

基金项目: 国家科技部重点研发计划中医药现代化重点专项资助(2023YFC3504101);药品监管科学全国重点实验室第二批课题、广东省药品监督管理局科技创新项目资助(2022TDB11)

作者简介: 王淑红,女,硕士,主任药师 研究方向:中药质量控制及安全性、有效性评价与中药智能化及数字化研究;李君瑶,女,硕士,主管药师 研究方向:中药检验及中药质量标准研究。王淑红和李君瑶为共同第一作者 * **通讯作者:** 王冰,男,博士,副主任药师 研究方向:药物安全性和有效性评价新方法研究 Tel:(0755)26031805;马双成,男,博士,研究员 研究方向:中药质量控制和评价研究 Tel:(010)67079501

准数字化水平不断提高”作为战略目标之一,提出发展机器可读标准、开源标准,推动标准化工作向数字化、网络化、智能化转型。同年12月,国家标准委、中央网信办、科技部等10部门联合印发《“十四五”推动高质量发展的国家标准体系建设规划》,提出深入推进国家标准数字化试点,探索增加机器可读标准、开源标准、数据库标准等新型国家标准供给形式;探索建立支撑国家标准数字化转型的信息系统。2022年12月,中共中央、国务院印发《扩大内需战略规划纲要(2022—2035年)》中提出要积极参与数字领域国际规则和标准制定。2024年3月,市场监管总局等18部门联合印发的《贯彻实施〈国家标准化发展纲要〉行动计划(2024—2025年)》中明确了要积极推进标准数字化研究,构建标准体系框架,开展标准数字化试点。

1.1.2 国外数字化政策^[3-6] 国际标准化组织(International Organization for Standardization, ISO)与国际电工委员会(International Electrotechnical Commission, IEC)提出了SMART标准(Standards Machine Applicable, Readable, and Transferable)的概念,旨在将标准文档转化为机器可读、可理解、可执行、可解释的文档,以推动标准化工作的数字化。欧盟通过《数字化转型战略计划》及后续战略,明确将标准数字化作为工业转型的关键,并启动多个项目促进标准数字化转型。美国则在《机器智能国家战略报告》中强调标准数字化的重要性,美国标准化协会亦将SMART标准纳入战略议题,推动数字化工具优化全球标准制定与传播。英国政府发布《第四次工业革命的标准:释放标准创新价值的HMG-NQI行动计划》,其中数字化转型行动聚焦于加速标准数字化,包括发展机器可读标准、改革数字平台提升可访问性,并开发相关框架与培训材料。德国则在其《标准化战略》中,将“机器可执行标准”视为核心驱动力,并发布白皮书详细规划至2030年的标准数字化路径,探索人机协同制定、处理及调整数字标准的未来图景。

1.2 数字技术现状^[6]

标准数字化分为通用技术方法和具体行业应用实践两个层面。通用技术方法层面关注于标准数字化中的一般性问题,提供普遍适用的数字化解决方案和指导。具体行业应用实践则着重于将这些通用技术方法结合行业特点实际应用,以提升准确率、效率和易用性。在不同的专业领域,这可能包括使用行业专用术语以提高标准文本处理

的准确性,或是在生产行业中根据产品属性建立数据字典和采集质量控制数据。这种实践推动了标准数字化的具体发展,尽管存在一些挑战,如技术通用性不足,难以总结推广,以及数据共享不足导致的数据孤岛问题。

无论是哪个层面,相关的信息技术如集成电路、通讯技术、区块链和图像处理等都是实现标准数字化的关键技术。这些技术支持了从文本处理到数据管理、从人机交互到安全保障的各个方面,确保标准数字化能在不同行业中有效实施。

1.3 应用现状^[6]

目前,标准数字化在很多国家、地区、行业、组织、企业中都在进行探索和实践。

1.3.1 ISO/IEC 联合开发的标准在线开发平台

ISO和IEC联合开发了面向全球数万名标准化专家的在线标准开发平台(online standards development, OSD)。目前,该平台可供工作组进行试用。该平台的设计目标是:从标准创作过程开始就更有组织和更高效地协作工作;在标准开发的早期阶段改进/简化内容质量;提高内容的整体质量;便于评论和最终汇编决议;基于标准的内容开发系统(The National Information Standards Organization Standards Tag Suite, NISO STS);统一ISO/IEC标准开发和流程。

1.3.2 中国标准化研究院国家标准馆标准服务系统 中国标准化研究院国家标准馆是我国唯一的国家级标准文献和标准化图书情报馆藏、研究和服务机构。标准馆利用自然语言处理等信息技术手段,对标准文献进行结构化加工,提取标准中关键技术指标,建成了覆盖国际主要国家、国内各类标准的文献资源库和信息库,开展了标准知识库构建、标准信息数据库构建、专题数据库建设、标准内容揭示和指标比对、标准大数据分析等相关研究和服务。

1.3.3 中国国家电网有限公司的制度标准智能管理系统建设^[7-8] 国家电网有限公司通过人工智能(AI)技术对企业标准和作业指导书等基础数据进行深度知识加工,以构建电网运营监测分析的知识图谱。这一转型旨在将标准识别、执行和反馈自动化,创建一个动态、响应迅速、按需使用、互动友好且智能辅助的信息系统生态。知识图谱的构建分为3个阶段:首先,利用电力词典创建电力词汇图谱;其次,结合专家和机器标注提取电力制度标准关键词及其关系,形成制度标准知识图谱;最后,整合设备、规划、运行、客户信息等,形成全面的电网运营监测分析知识图谱,以提升电网运营的智能化水平。

2 中药标准的数字化转型升级

中医药高质量发展的关键核心是建立科学合理的中药质量评价体系,只有不断创新中药质量评价模式,才能有效服务产业高质量发展。因此,中药标准的数字化转型成为催生新质生产力,服务国家重大战略“中医药现代化”的关键环节。

2.1 目前中药标准的局限性

2.1.1 标准呈现形式 中药标准体系构成复杂、层级多,目前主要以纸质文本或电子版 PDF 格式发布。这种形式受到载体本身的限制,存在局限性:

信息孤立和共享困难。纸质和 PDF 格式难以实现数据的统一整合和共享。各标准之间信息孤立,缺乏关联性,使用时难以通过简单的搜索功能找到跨文档、跨标准的相关信息,限制了标准的综合利用和关联分析。

更新和维护不便。中药标准需要定期更新,现在的呈现形式更新过程烦琐且耗时,更新后的版本需要重新分发,增加了成本。

智能分析和应用受限。现代信息技术的发展使智能分析和大数据处理成为可能,而这些技术难以直接应用于中药标准的呈现形式。无法进行自动化的数据挖掘、智能解析,限制了中药标准在研发、生产和监管中高效应用。

环境问题。大量的纸质文档不仅增加了印刷和存储成本,还对环境造成不小的负担。

2.1.2 质量控制模式 现有的中药质量控制模式多采用传统经验鉴别与化学成分分析相结合的方式^[9]。传统的经验鉴别即通过对中药的外观性状、颜色、气味等形态特征进行初步评价^[10]。在化学成分分析过程中,采用了光谱、色谱等分析仪器进行数据采集和结果分析。随着新兴科学技术的迅猛发展,传统的经验鉴别开始融入仿生学、AI 等技术,产生了如机器视觉、嗅觉、味觉技术。这些技术使得中药感官特性得以量化。这种现代科学技术和传统检验鉴别方法的结合,给中药质量控制模式提供了更多的可能。当前,传统的经验鉴别仍然作为主要评价方式。这种方式高度依赖检验人员的经验,容易受到主观因素的影响,导致不同人员可能得出不同的结论。特别是对于经验不足的新手,误判的风险相对较高。AI 等新兴科学技术在中药质量控制中的应用还处于探索的阶段,要使这些技术真正服务于中药的科学监管,需要建立与其适配的平台。

2.1.3 标准物质使用 中药标准中的薄层鉴别、

特征图谱、指纹图谱、含量测定以及安全性检查等项目需要使用大量的实体标准物质,这大大增加了检验成本。这些标准物质有的提纯困难且价格高昂,有的不稳定、毒性大且存在安全风险,有的目前已经受到管制无法购买,甚至停产。Wang 等^[11]总结了近些年发展出的多种标准物质的形式及优缺点,发现目前的数据库智能化程度不高,多数为碎片化的孤立数据库,尽管产生了如一测多评法、双标线性校正法、对照提取物法等新技术、新方法,但就目前来说,仍然未能完全摆脱中药质量控制依赖于实体标准物质大量使用的格局。

2.1.4 数据资源应用 自动化分析仪器如光谱、色谱技术的不断普及,能够快速、精准地生成海量数据,大幅提高了中药质量检测的效率和准确性。然而,这些数据往往分散在不同的仪器和系统中,缺乏统一的标准和平台,无法实现互联互通。数据共享和整合困难,大量有价值的数据未能充分挖掘和利用,造成数据资源浪费。

2.1.5 标准管理工作模式 标准化工作的全生命周期涉及对立项、起草、征求意见、审查、实施、修订、废止等过程的管理。现阶段中药质量控制领域已建立较为完善的标准体系,目前标准管理工作主要依赖人工,尚未构建数字化的标准全生命周期管理系统,存在以下局限:标准体系更新速度慢、标准制修订协调过程长、决策层面的信息通报周期长,较大程度上影响了标准时效性;管理工作需要大量的时间和资源,人工编辑和审查文本需要耗费大量的人力,耗时长、工作量大,且过程中易出现错判、遗漏的情况。

综上所述,这些问题凸显了中药标准数字化转型的迫切需求。数字化和智能化技术的快速发展为这些问题提供了解决思路。利用卷积神经网络等图像识别技术已经可以实现高精度的物种识别^[12];计算机视觉技术提取中药饮片横切面图像纹理特征参数已在中药材鉴定方面获得成功应用^[13-14];中药数字标准物质(DRS)替代实物标准物质以及中药指纹图谱相似度在药品定性和定量分析中的应用研究^[9]。这些研究成果为中药标准的数字化转型奠定了研究基础。

2.2 中药标准数字化转型

当前,国内对标准数字化概念认知有一些基本的共识。标准数字化转型一是标准文档利用数字技术改变标准本身形态,二是标准化过程。根据 IEC 与 ISO 关于标准数字化转型的说法,机器实现自主

理解和处理标准的前提是标准形式与数据结构的数字化,使标准以数字形式展现,从而由机器自动识别,无须人员参与,实现检测与标准的自动化。中药标准的数字化转型具体包括两个方面:中药标准管理数字化和中药数字标准。

中药标准管理数字化即利用信息化、数字化手段对标准全生命周期进行数字化管理维护。标准制订和发布形式的变更是中药标准管理数字化的重要体现。借助标准数字化转型的基础工作,能够实现中药标准在线编写、意见征集、在线审议、检索、修订、更新和废止等,协调标准体系动态更新与标准制修订过程的同时,加快与决策部门之间的信息流动,最终实现缩短标准的制修订周期,提升标准编制效率的目的;此外,还可以根据标准提高要求动态反馈标准更新过程,得到的海量数据信息为标准实施应用提供起草、复核原始数据支撑(图1)。

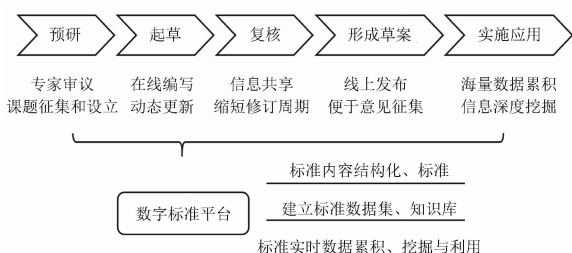


图1 中药标准管理数字化的研究思路示意图

中药数字标准是指借助一系列数字化、智能化技术,突破中药标准的传统方式和传统路径,将评价中药真伪优劣的法定方法转化成为由机器自动读取、分析和结果判断的数据形态。目前,中药数字标准属于理论与技术创新研究阶段。

2.3 中药数字标准的意义

中药数字标准的开发和应用对于中药行业的高质量发展具有重要意义。它不是忽视、放弃传统的中药标准,而是应用新技术改造传统方式,通过技术革命性突破,实现中药行业标准的智能化、绿色化,高效服务中药产业高质量发展(图2)。

2.3.1 加强标准引领和质量支撑 中药数字标准的开发与应用,是将数字科技融入中药质量决策过程的一次重大创新。通过数字化、智能化的手段,保证中药控制过程更加科学、客观。这种创新不仅体现了中药监管科学对于高科技、高效能、高质量的要求,更展现了中药数字标准在数字化、智能化、客观化、低碳化等方面的先进生产力质态。这种质态的

转变,加强了标准引领和质量支撑,打通了束缚中药高质量发展的堵点和卡点。

2.3.2 强中药行业的创新发展 传统的中药研发过程往往依赖人工操作和经验,效率低且难以保证质量的一致性。引入数字标准后,中药质量控制的每一个环节都得以实现数字化管理,极大地提高了工作效率,保证药品质量控制的稳定可控性。推动传统中药高端化、绿色化、智能化转型,推动了中药行业的创新发展。

2.3.3 提升中药国际竞争力 在全球化背景下,药品的国际竞争日益激烈。数字标准的引入不仅使中药行业能够与国际标准接轨,提升产品的国际市场竞争力,还有助于中药文化的传承和全球推广,助力提升中药的国际竞争力。

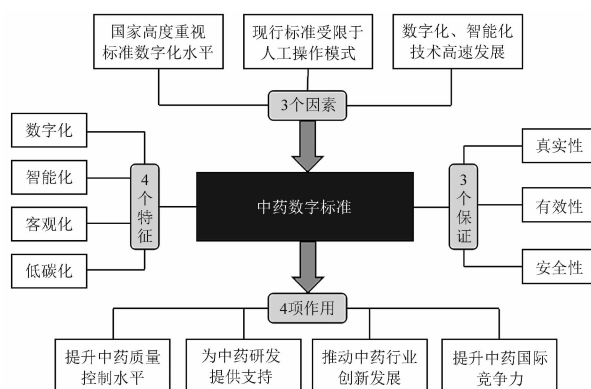


图2 中药数字标准逻辑关系示意图

3 中药数字标准概念的内涵和外延

中药标准数字化是利用数字技术和智能设备,对中药标准及相关数据进行加工、处理、解析、标注和关联,使其能够被机器自动识读、理解、调用和访问,满足不同场景的应用。同时,这一过程还包括对标准全生命周期(立项、起草、制修订、征求意见、审查、复审、批准、应用、实施、废止等)的管理。中药标准数字化,其核心要素就是中药数字标准。

“中药数字标准(Traditional Chinese Medicine Digital Standards, TCM-DS)”,是一种无人工干预的机器可读、可执行、可解析、可决策标准。该标准借助数字技术(大数据、AI、云计算等)、智能控制技术(智能信息处理、智能信息反馈和智能控制决策),将中药现行法定标准转化成为机器可识别的数字模型(图像、公式等),并通过机器进行自动读取、在线分析和结果判定(图3)。

3.1 中药数字标准的组成

中药数字标准由数字模型、机器两部分组成。



图3 中药数字标准示意图

数字模型的建立基于图片、图谱、数值等数据信息。通过对数据的采集、预处理、特征提取建立相应模型,并开展后续的训练、优化、验证,得到具有良好泛化能力的模型,确保模型用于待测样品的分析、结果判定时的准确度和稳定性。数字模型的建立是基于具有代表性的、来源准确的样品,模型的建立不仅需要关注数据量的大小,还需要综合考量模型的复杂度、假设、数据表示方式等多个方面。

机器是按特定指令执行标准任务并给出决策支持的智能设备与分析软件的总称。智能检验设备以多种先进的传感器技术为基础,在计算机系统控制下,通过自动控制技术,数字技术自动实现对待测样品的数据(图像、图谱、数值等)采集和特征识别。智能检验设备按预定程序自动完成工作任务,在这过程中,人员只需设定程序无须直接参与,在保证检验结果准确性的同时,大幅提高检验效率。分析软件由多个基本程序组成,可提供人机对话环境并实现机器自动分析。它具有开放性的程序结构和多种接口,包括模型数据输入和输出,分析参数的设置等,分析软件可对输入的数据信息进行识别与预处理,在数据库内开展匹配检索并对检索结果进行整合分析,提供准确、可靠的决策支持。分析软件可以根据需求进行二次开发,以匹配不断变化的分析要求。

3.2 中药数字标准的特征

机器可读。将评价中药真伪优劣的法定方法转化成为机器可识别的数字模型(模型、公式等),并通过机器进行自动读取、在线分析和结果判定。

执行客观。中药数字标准建立与执行的先决条件是信息化,经由信息系统将方法、流程固化下来,标准实施全过程由计算机监督执行;标准内容被转化为数字模型,直接由机器读取,无人为主观因素干扰,保证检验结果的客观性。

动态更新。可依据检验数据或研究结果申请增删或修订标准中存在缺陷的项目,实现中药数字标准起草、复核、上报、审议、发布、废止这一全生命周期的数字化管理,实现标准的快速迭代与动态更新,确保其适用性与适用性,减少标准制修订人工成本,缩短中药标准制修订周期。

交互智能。中药数字标准具有开放性的程序结构和多种接口,实时提供人机对话环境,可实现自主应答询问、提供决策支持智能转型。

低碳环保。中药数字标准采用DRS替代传统标准物质,可大幅降低实物标准物质导致的资源浪费,使数据信息以低碳、环保的方式传递和共享。

4 中药标准数字化的实现方法

4.1 数字技术

在中药标准的数字化转型过程中,AI技术发挥了至关重要的作用。AI技术在图像识别分析、数据挖掘处理、模型预测及决策支持方面具有显著优势,为中药材(饮片)鉴别^[12]、成分分析^[15]等方面的研究提供了新的手段。AI的核心研究领域是机器学习,深度学习是一种深层的机器学习模型,其本质是一种模拟人脑神经系统处理复杂信息的数学模型和计算结构系统的人工神经网络^[16-18](artificial neural network, ANN)。ANN广泛应用于计算机视觉、自然语言处理、图像处理等领域,为中药性状、显微鉴别以及薄层色谱项目提供了客观化、科学化的标准。

目前,ANN的主流模型是卷积神经网络(convolutional neural networks, CNN),通常包含卷积层、池化层、全连接层、输出层,在图像识别和处理领域得到了广泛的应用^[19]。CNN可实现对中药饮片特征参数的提取与分析,为相近品种、易混淆品种和不同产地中药饮片鉴别提供可视分析手段。Sun等^[20]运用CNN在测试的50种中药饮片图像中可以实现70%的平均识别精度。Chen等^[21]基于卷积神经网络建立了一种无损快速识别陈皮及广陈皮的方法,识别准确率达到94%。CNN可用于中药性状、显微鉴别。

由于神经网络具有从复杂数据中寻找规律的特性,还被广泛应用于中药模式识别领域。Hu等^[22]采用高效液相色谱法指纹图谱技术结合神经网络鉴别陈皮药材的产地与种源,在对新会陈皮和非新会陈皮的识别训练与预测中识别准确率达到100%。Sun等^[23]采集9种不同来源丹参饮片的高光谱,结合神经网络建立了分类模型,7种丹参样品判别结果的真正率、命中率和特异度均达

到 100% ,其余 2 种丹参样品的真正率、命中率和特异度也不小于 90% 。

多层感知器(multilayer perceptron, MLP)是一种前馈人工神经网络模型^[24],可以更加精准地分析和判断样本的成分和性质,具有良好的自学习和自适应能力。MLP 可适用于中药薄层鉴别项目。MLP 能够对薄层色谱图像中的特征斑点进行精确分析和量化处理,细致地区分斑点的大小、颜色和形状等特征;通过算法,进而提取出特征斑点的颜色和距离信息,通过自监督学习,逐渐逼近真实的映射关系,从而形成一个准确的计算公式,MLP 会根据自己学习到的映射关系给出预测结果。

4.2 数字模型

数字模型的建立是中药数字标准的核心内容。数字模型包括两类:一类是图像模型,通过 AI 视觉识别技术对图像数据进行采集,提取特征而得,如性状、显微鉴别项。另一类为公式模型,以具体数学形式表达。根据变量的复杂程度,可分为简单公式模型和复杂公式模型。简单公式模型只包含少量的变量和操作,结构简单,模型的核心就是数学公式。复杂公式模型在简单公式模型基础上,变量的数量和操作层次都不同程度增加,模型结构更为复杂(图 4)。

以采用图像模型的性状、显微鉴别项,简单公式模型的水分项,复杂公式模型的薄层色谱、特征图谱、指纹图谱、含量测定项为例,简单阐述各类数字模型的构建思路。

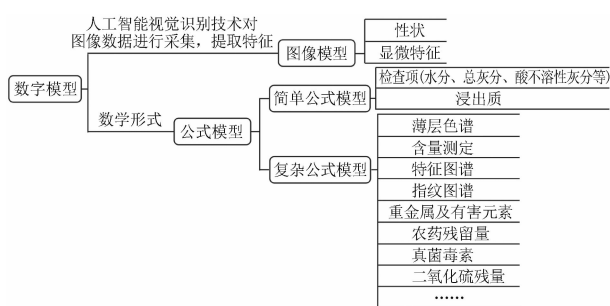


图 4 数字模型分类示意图

4.2.1 图像模型应用——性状 数据采集。性状图像模型需要采集中药材(饮片)外观特征数据。收集基原准确的中药材,涵盖不同主产地、不同采收期、不同加工方式等,建立图像标准化采集流程。为获得稳定的图像数据,可以采用统一的图像采集设备(图 5A),从多视角收集样品图像数据,并对药材图像进行自动定位和提取。按一定的比例将样品划

分得到训练集和试验集,分别用于中药数字标准性状模式识别模型的建立和测试。

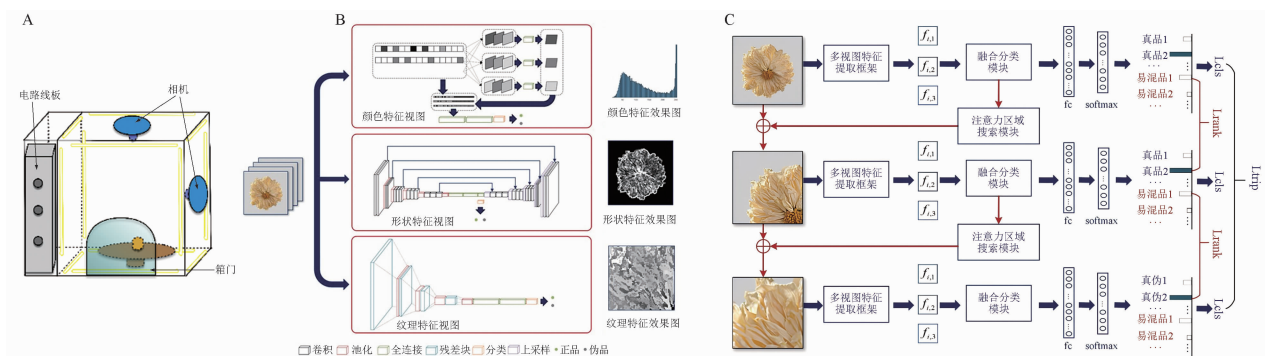
特征提取。将中药材(饮片)颜色、形状、表面等外观特征图像数据,分别引入关系网络、U-net 网络和残差网络进行特征提取和学习记忆。其中,关系网络通过学习和挖掘图像中不同特征之间的关系,提取出中药材的颜色和形状等全局特征。这种网络结构可以有效地捕捉图像中的复杂模式,并对其进行准确分类。U-net 网络则专注于提取中药材表面的局部细节特征,在保留图像细节的同时,对图像进行逐层抽象和特征提取,这种网络结构对于中药材表面纹理和质地的识别具有很好的效果。残差网络则用于提取中药材图像中的深层次特征。通过引入残差结构,网络能够更好地学习图像中的复杂信息和模式,从而提高分类的准确性。

每种网络都独立进行特征提取和学习记忆,提取结果通过融合算法进行融合输出。这种策略可以充分利用不同网络的优势,提高分类的准确性和泛化能力。构建多视图细粒度分类框架,实现特征学习互不干扰,提取结果融合输出(图 5B)。

模型建立。数字模型由多层子网络构成(图 5C)。数字模型的多层子网络结构类似于人类视觉系统的层次结构,从低级的边缘、纹理特征,到高级的形状特征,层层递进。每一层子网络都负责提取和抽象输入数据的特定特征,然后将这些特征传递给下一层子网络。这种层次化的结构使得模型能够逐步从原始数据中提取出更加抽象、具有判别性的特征。

通过递归学习的方式,从粗到细地产生判别性区域。在每一轮学习中,模型都会对输入数据进行分类,并根据分类结果调整网络参数。随着学习的深入,模型能够逐渐关注到更加细微的特征,从而提高分类的准确性。这种递归学习方式使得模型能够自适应地调整其判别性区域,以适应不同的数据集和任务。利用多层次特征集成的优势,有效地整合不同子网络的特征分类结果。每层子网络都会提取出输入数据的不同特征,这些特征在分类过程中都具有重要的作用。通过将这些特征进行集成,模型能够综合利用各种信息,从而提高分类的准确性和鲁棒性。

通过上述的多层子网络结构、递归学习方式,利用多层次特征集成的优势,有效地整合不同子网络的特征分类结果,实现最终的细粒度分类。



A - 图像采集设备; B - 图像特征提取; C - 图像模型建立。

图5 性状图像模型建立示意图

4.2.2 图像模型应用——显微鉴别 数据采集。显微鉴别模型需要采集中药材(饮片)的显微组织特征数据,包括特征分类信息和形态分类信息。以菊花为例,显微组织特征包括5种,分别为T形毛、腺毛、花粉粒、草酸钙以及草酸钙(偏光),每种显微特征的形态又分为典型和非典型两种(图6A)。

特征提取。采用层级分类的方式,首先使用“预训练+迁移学习微调”的方法进行分类训练。预训练阶段,模型在大量未标注数据上进行无监督学习,以学习通用的特征表示。通过在大规模数据集上的预训练,模型可以捕获到图像中的基本结构和纹理信息。然后将预训练得到的模型迁移到具体的显微结构图像分类任务中,利用有标注的数据进行微调。这个过程可以使得模型更加专注于显微结构图像中的特定特征,从而提高分类的准确性。

在确定分类算法的大致框架后,采用自监督旋转不变性注意力的神经网络和注意力机制的卷积网络来进一步提取显微结构图像的特征。自监督学习是一种利用未标注数据进行训练的方法,它可以帮助模型学习到更多的图像结构和纹理信息。旋转不变性注意力机制则可以帮助模型更好地应对图像旋转等变化,提高模型的鲁棒性。而注意力机制的卷积网络则可以在特征提取过程中,将更多的注意力集中在图像的关键区域,从而实现了对显微结构图像典型和非典型特征的自动提取。

具体来说,显微结构图像输入到自监督旋转不变性注意力的神经网络后,通过自监督学习的方式,让模型学习到图像中的基本结构和纹理信息。将提取到的特征输入到注意力机制的卷积网络中,更多的关注力集中在图像的关键区域。通过这种方式,可以实现对显微结构图像典型和非典型特征的自动提取,从而实现了对显微结构图像的有效分类(图6B)。

模型建立。以菊花的显微组织为例,首先需要

训练一个五类分类模型,这些类别包括T形毛、腺毛、花粉粒、草酸钙以及草酸钙(偏光)。每一类都具有其独特的形态特征和结构,这些特征是模型进行区分的关键。

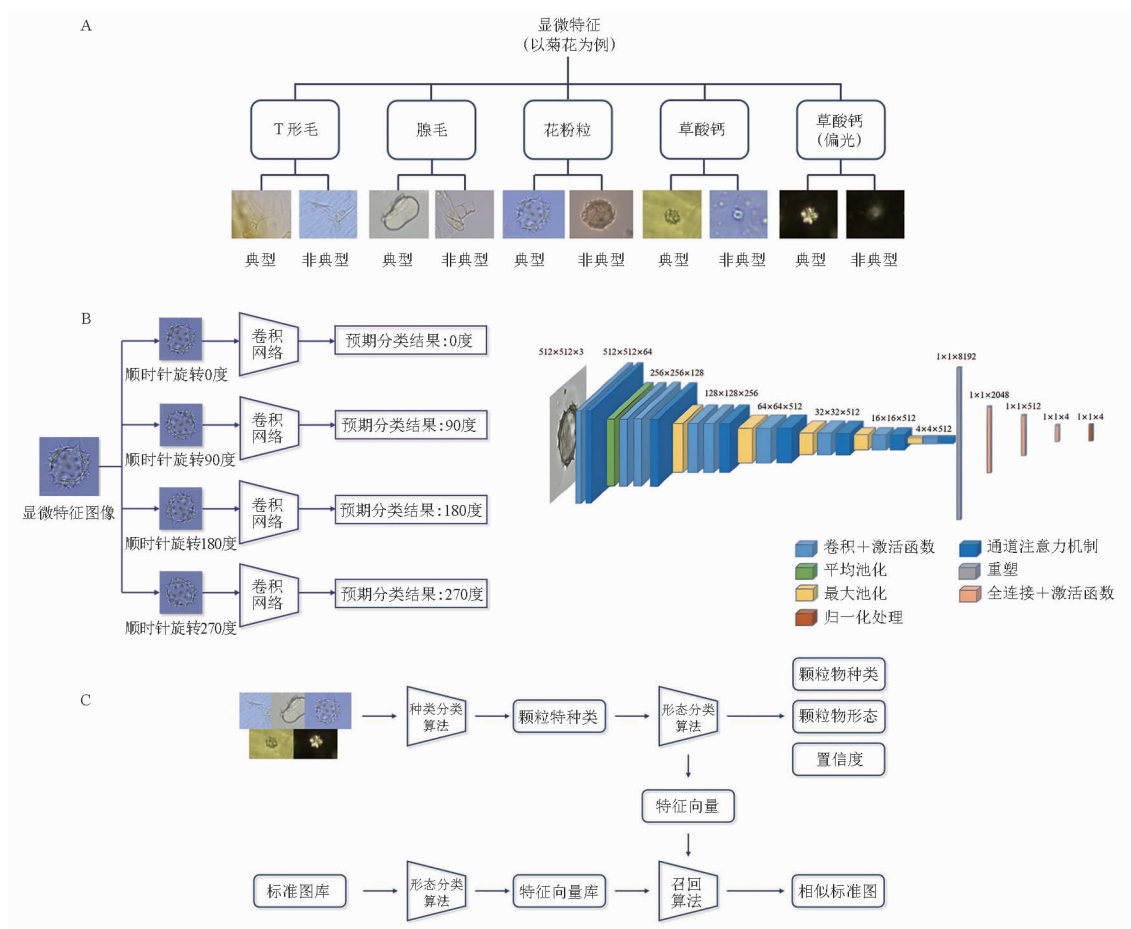
在训练过程中,还需要考虑典型和非典型样本的二分类问题。典型样本是指那些具有明显特征,容易区分的样本;而非典型样本则可能因为形态变化、模糊等原因而难以归类。将样本分为典型和非典型两类,分别进行训练,可以提高模型的准确性和鲁棒性。

在确定了显微组织特征种类、形态和判断置信度后,通过召回算法(一种基于特征相似度的搜索算法),根据输入图像的特征搜索并快速找到与输入图像相似的标准图,从而为研究人员提供参考和对比(图6C)。

4.2.3 简单公式模型应用——水分 中药水分测定是基于原理和实验方法进行的统计分析,核心在于建立一个可靠的公式模型,以样品的质量为变量,通过水分公式,能够准确计算出中药中的水分含量结果。这个公式的推导和验证,离不开大量实验数据的支撑和统计分析方法的运用。通过收集不同种类的中药样品,测定其水分含量,并将这些数据用于模型的构建和优化。

将水分公式及限度规定嵌入到自动化设备中,由机器自动识别模型并进行水分的自动计算与结果判定。这不仅大大提高了测定效率,还减少了人为误差,保证了测定结果的准确性和可靠性。

4.2.4 复杂公式模型应用——薄层色谱 薄层色谱模型需要采集薄层色谱斑点的距离、颜色数据信息(图7A)。通过对薄层色谱特征斑点的颜色和距离信息进行提取和量化,结合先进的机器学习方法,如MLP,可以更加精准地分析和判断样本的成分和性质(图7B)。



A - 显微组织特征分类(以菊花为例); B - 自监督旋转不变性注意力的神经网络和注意力机制的卷积网络示意图; C - 显微鉴别模型流程。

图 6 显微鉴别图像模型建立示意图

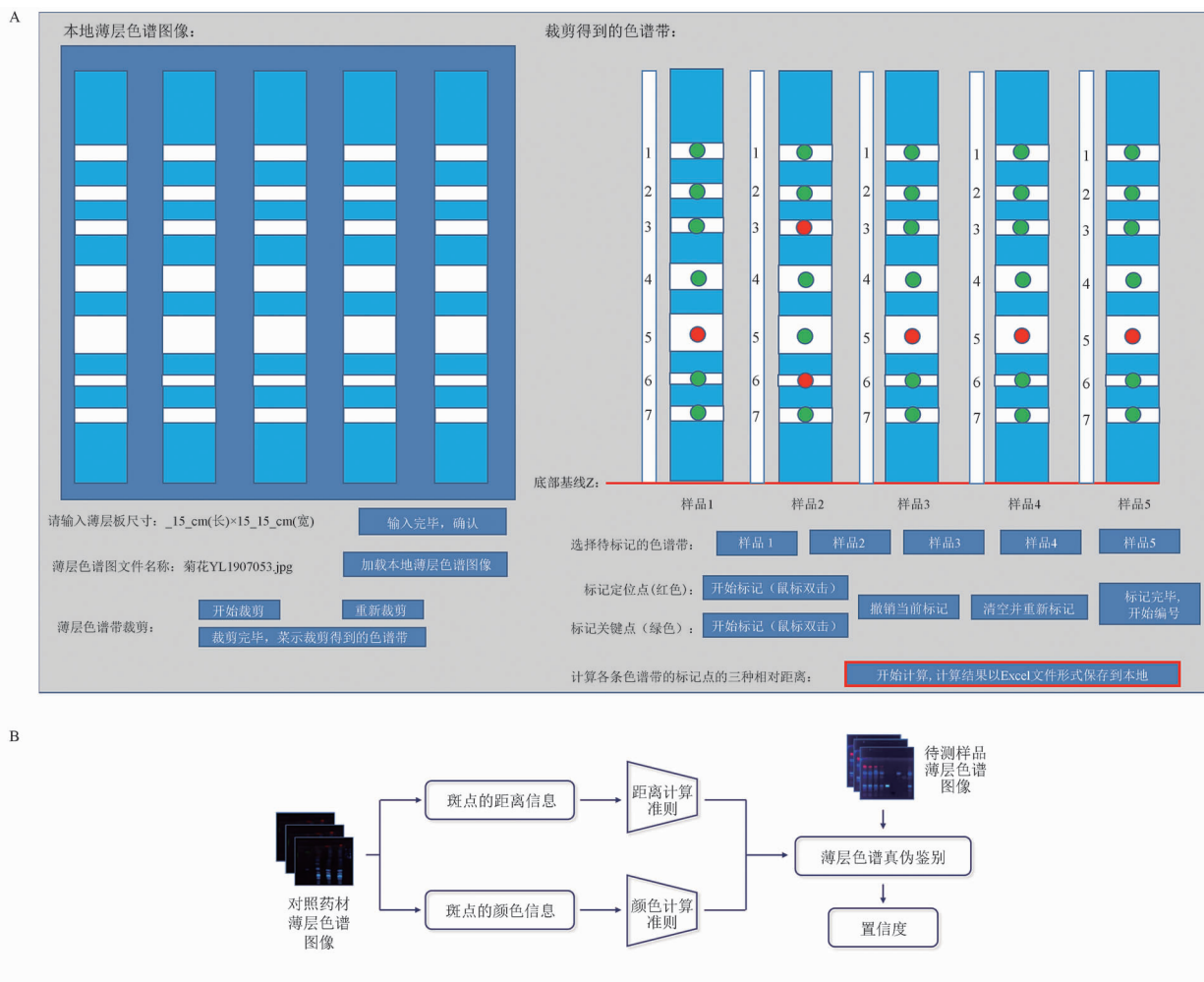
对薄层色谱特征斑点进行准确的提取和量化过程涉及到对色谱图像中的颜色、形状、大小等特征进行精细的分析和处理。通过专业的图像处理软件或算法,将色谱图像转化为数字化的信息,进而提取出特征斑点的颜色和距离信息,这些信息是后续机器学习模型训练的基础。

对上述得到的信息进行基于 MLP 的有监督学习。MLP 是一种常见的人工神经网络模型,具有良好的自学习和自适应能力。训练过程中,需要已知结果的样本数据作为训练集,让 MLP 学习如何从输入的特征信息中预测出相应的结果。通过不断调整网络参数,使 MLP 逐渐逼近真实的映射关系,从而形成一个准确的计算公式。

MLP 模型训练成功后,进行待测样本薄层色谱图像的采集,对其进行预处理和特征提取。将这些特征信息输入到训练好的 MLP 模型中,模型会根据已学习到的映射关系给出预测结果。同时,还可以结合少量的对照药材薄层色谱图像进行辅助分析,以提高结果的准确性和可靠性。

4.2.5 复杂公式模型应用 采用复杂公式模型的项目需要采用实物标准物质作对照,如中药质量控制特征图谱、含量测定、安全性控制等项目,随着互联网+、大数据等技术的快速发展及应用,标准物质的形式从实物标准物质、纸质图谱集等传统形式发展到替代标准物质、DRS 等创新形式。将标准物质数字化,既能减少实物标准物质的制备和标定,减少资源浪费,节省成本;又能以系统化、规范化的形式提供更全面的、相互关联的信息,达到智能搜索、大数据存储及互联网共享的目的。

特征图谱、指纹图谱、含量测定。中药指纹(特征)图谱是基于中药整体性控制思路出发,多采用色谱技术对其含有的化学成分进行系统的、整体的、专属的表征^[9]。现有标准操作方法表现为:特征图谱与对照品保留时间进行整体比较,从而完成对多个成分的鉴别;指纹图谱是将待测样品的测定图谱与对照品或对照品图谱进行整体比较,对多批次样品色谱图进行共有峰的提取与标定,对非共有峰则规定数量及峰面积限度,采用夹角余弦法、相关



A - 薄层色谱信息提取; B - 薄层色谱模型流程。

图7 薄层色谱模型建立示意图

系数法、相对熵法、峰重叠率法、总量统计矩相似度法等相似度算法进行量化。经对比验证,《中国药典》2020年版配套软件——中药指纹图谱计算机辅助相似度评价软件采用夹角余弦法计算色谱峰的相似度。

《中国药典》2020年版收载指纹(特征)图谱的数量增加到73个,并配套中药色谱指纹图谱相似度评价系统软件,实现色谱数据预处理、峰匹配以及相似度计算。中药色谱指纹图谱相似度评价系统软件的应用,是中药标准数字化的雏形。Sun等^[25]开展了数字化指纹图谱研究,利用数学原理或方法对指纹图谱本质特征进行定义、表达、重组、模拟,将检测方法和仪器检测的信号数据转化为容易鉴别且能揭示中药复杂信息特征的数字化信息。

中药安全性评价项目。引起中药安全性问题的物质主要包括外源性污染有毒有害物质和内源性有毒物质,中药安全性评价项目重点关注的是残留量

值不得超过或不得检出,需要根据限度要求建立科学、合理的中药外源性污染有毒有害物质和内源性有毒物质的含量限度模型。

4.2.6 中药数字标准物质 标准物质是标准执行的重要保证,也是实物化的标准。DRS也是数字化标准的重要组成部分,是实物标准物质的数字化、数据化、互联网化及智能化形式,通过不同的维度反映实物标准物质的物理、化学和生物等特征。DRS不仅是标准物质自身形式的数字化,同时也是包含仪器性能、色谱柱等影响因素的融合信息数据库(DRS Database)^[11]。通过相似度检索进行跨库联合匹配,采用“双线性校正法”“保留时间校正法”等方法预测保留时间来对样品图谱进行定性鉴别,采用基于大数据的色谱柱正负列表,实现实物标准物质的数字化代替。

中药DRS模型。中药DRS模型是在DRS量值稳定以及量值传递快速、高效的基础上发展起来的,

将实物标准物质的属性、性能等参数标准化,并形成

数字形式的计量标准用于定性分析或定量测定(图8)。

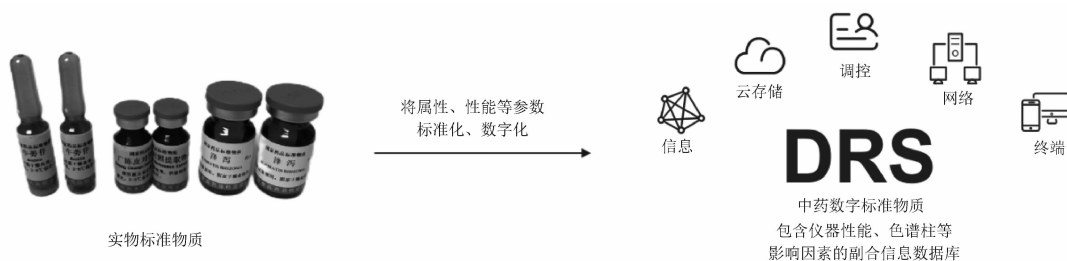


图8 数字标准物质模型的形成示意图

DRS 模型应用。以 DRS 为基础的复杂公式模型应用包括特征图谱和含量测定。通过 DRS 模型,可以对待测样品进行保留时间的准确预测,并确定系列特征成分峰的位置;对待测样品含量结果数据进行运算分析,在无实物标准物质的情况下实现准确定量,这将解决实物标准物质生产、储存过程中导致的资源浪费和环境污染问题。

5 中药数字标准面临的问题与挑战

在以信息技术为代表的新一轮科技革命和产业变革的加速推进的当下,中药标准的数字化进程势不可挡。随着中药数字标准的概念、技术途径和体系架构的不断成熟,相应地,各种问题和挑战也随之浮现。

5.1 中药数字标准面临的问题

中药材来源复杂,成分差异较大,如何收集足够数量来源可靠,具有广泛代表性的样品,是建立可靠模型的基础和挑战。

中药标准数字化涉及一系列当下最先进的 AI 技术,包括数字化加工、自然语言处理,知识图谱构建等技术,如何攻破上述技术难题是中药数字标准转型的关键。

中药标准数字化涉及国家标准、行业标准、团体标准、企业标准以及企业内控标准,存量数据多。而标准数字化对离散化程度要求高,加工难度大,需根据重要程度对上述标准进行差异化加工标引,打好数据基础,从而保证将来的应用。

为确保中药标准数字化建设能有效服务于药品监管和产业发展,需要针对不同的应用场景实现智能对接,并通过实际应用不断提升中药数字标准的适用性。

中药数字标准的研究成果最终需与国家数字化标准体系对接。当前,国家层面的总体设计与技术规范仍在制定过程中;同时,中药数字标准的法律地

位尚需时间来确立和完善。

5.2 中药数字标准面临的挑战

管理的挑战。标准数字化不仅涉及技术工作,还需要开展相应的管理改革,重新审定标准制修订流程与审批管理权限,确保各项工作能够高效、有序地进行。数字化并不意味着完全摒弃人为干预,相反,在适当的时候,人的参与和判断仍然是不可或缺的。因此,需要在管理上明确何时以及如何介入,确保标准数字化的顺利推进。

AI 技术的挑战。AI 技术在大数据分析中的应用已经取得了显著的成果,但在实际应用中,也必须正视其存在的问题。例如,AI 的分析结果仍存在结果可解析性差的问题,且无法阐明结果的生成逻辑;在一些重要的应用场景中,AI 输出结果的可信性仍然值得探讨。标准数字化采用 AI 技术来理解语义,同样也面临上述困境。必须关注标准数字化过程中的责任与伦理问题,确保 AI 技术的健康发展。

网络安全的挑战。标准数字化高度依赖于计算机系统的应用,这使得黑客和木马攻击等网络安全问题可能会给标准数字化工作带来巨大风险。因此,需要高度重视其关联信息系统的安全性,重要标准不能完全依赖机器解读,仍需要人为把关。

6 结语与展望

随着科技的迅猛发展,中药标准的数字化已成为提升中药质量控制,推动产业高质量发展的必然趋势。国家主管部门应明确中药标准数字化的战略意义,并制定详尽的发展规划,明确目标、任务和时间表。在此基础上,出台相应的政策法规,为中药标准数字化提供坚实的制度保障和明确的法律依据。同时,加大资金投入,支持相关科研项目和基础设施建设,特别是构建国家级别的信息化技术平台助力中药标准的数字化转型,为其提供庞大数据信息存储空间

传统中药标准的数字化新形态

王淑红^{1,2}, 李君瑶^{1,2}, 郭子瑜^{1,2}, 王子骏³, 何轶^{4,5}, 苏畅^{1,2}, 谷丽菲^{1,2}, 卢光明³, 王冰^{1,2*}, 马双成^{4,5*}

(1. 深圳市药品检验研究院, 广东 深圳 518057; 2. 国家药品监督管理局中药质量研究与评价重点实验室, 广东 深圳 518057; 3. 哈尔滨工业大学(深圳), 广东 深圳 518055; 4. 国家药典委员会, 北京 100061; 5. 药品监管科学全国重点实验室, 北京 100061)

摘要:目的 数字时代下,中药标准数字化转型成为提升中药质量控制,推动产业高质量发展的必然趋势。**方法** 笔者研究团队利用 AI 技术在图像识别分析、数据挖掘处理、模型预测及决策支持等方面的优势,开展传统中药标准的数字化转型。**结果与结论** 尝试构建一种无人工干预的机器可读、可执行、可解析、可决策的中药标准数字化新形态,积极推动中药标准的数字化、网络化、智能化发展,服务标准数字化水平不断提高的国家战略目标。

关键词: 中药数字标准; 中药标准; 数字化转型; 新质生产力; 数字技术

doi:10.11669/cpj.2024.24.013 中图分类号:R282 文献标志码:A 文章编号:1001-2494(2024)24-2396-11

New Digital Form of Traditional Chinese Medicine Standard

WANG Shuhong^{1,2}, LI Junyao^{1,2}, GUO Ziyu^{1,2}, WANG Zijun³, HE Yi^{4,5}, SU Chang^{1,2}, GU Lifei^{1,2}, LU Guangming³, WANG Bing^{1,2*}, MA Shuangcheng^{4,5*} (1. Shenzhen Institute for Drug Control, Shenzhen 518057, China; 2. NMPA Key Laboratory for Quality Research and Evaluation of Traditional Chinese Medicine, Shenzhen 518057, China; 3. Harbin Institute of Technology, Shenzhen 518055, China; 4. Chinese Pharmacopoeia Commission, Beijing 100061, China; 5. State Key Laboratory of Drug Regulatory Science, Beijing 100061, China)

ABSTRACT:OBJECTIVE In the digital era, the digital transformation of traditional Chinese medicine standards has become an inevitable trend to improve the quality control of traditional Chinese medicine and promote the high-quality development of the industry.

METHODS The research team has leveraged AI technology's advantages in image recognition and analysis, data mining and processing, model prediction, and decision support to carry out the digital transformation of traditional Chinese medicine standards.

RESULTS AND CONCLUSION They aim to construct a new form of traditional Chinese medicine digital standards that is machine-readable, executable, interpretable, and decision-capable without human intervention. This effort actively promotes the digitalization, networking, and intelligent development of traditional Chinese medicine standards, contributing to the national strategic goal of continuously improving the level of standard digitization.

KEY WORDS: traditional Chinese medicine digital standard; traditional Chinese medicine standard; digital transformation; new quality productive force; digital technology

中药标准作为中药质量控制方法的重要载体,是药品监管的重要工具。2023年,国家药品监督管理局关于《进一步加强中药科学监管促进中药传承创新发展若干措施》(国药监药注〔2023〕1号)要求“提升中药标准数字化管理水平。建立完善中药国家药品标准、药品注册标准动态数据库,加快推进数字化标准建设,及时更新数据,实现药品标准的发布、查询、分析、研究、维护信息化^[1]。”中药标准迈向数字化转型新阶段。

1 全球标准数字化进展

标准是经济活动和社会发展的技术支撑,是国家基础性制度的重要方面^[2]。近年来,在全球范围内,国际标准化组织和欧洲、美国及亚洲的主要经济体均围绕标准数字化及相关主题开展研究和实践。

1.1 数字化政策现状

1.1.1 国家数字化政策^[2-3] 2021年10月,中共中央、国务院印发了《国家标准化发展纲要》,将“标

基金项目: 国家科技部重点研发计划中医药现代化重点专项资助(2023YFC3504101);药品监管科学全国重点实验室第二批课题、广东省药品监督管理局科技创新项目资助(2022TDB11)

作者简介: 王淑红,女,硕士,主任药师 研究方向:中药质量控制及安全性、有效性评价与中药智能化及数字化研究;李君瑶,女,硕士,主管药师 研究方向:中药检验及中药质量标准研究。王淑红和李君瑶为共同第一作者 * **通讯作者:** 王冰,男,博士,副主任药师 研究方向:药物安全性和有效性评价新方法研究 Tel:(0755)26031805;马双成,男,博士,研究员 研究方向:中药质量控制和评价研究 Tel:(010)67079501