

试论大数据对我国中药资源数据整合与应用的启示

田虹 田侃

(南京中医药大学, 江苏南京 210023)

摘要 大数据是信息技术产业又一次重要的技术变革,将为人类社会的发展带来不可估量的变化。大数据追求覆盖整体信息的宏观理念与中药资源可持续发展的概念具有一定程度的相似性,有必要推动大数据理念及其技术在中药资源领域的合理应用。中药资源数据主要包括我国中药资源种类、分布、蕴藏量、资源变化趋势、栽培与野生情况、收购量、需要量、质量等,这些数据对中药资源的调研、普查、评控、种植、流通、可持续化发展等方面具有重要意义。而当前由于中药资源管理主体存在差别、中药资源人才缺乏和潜在的安全威胁,对于这些数据的采集、存储、使用均存在一定问题,应进一步转变思想观念、提高技术水平、完善法律制度,将大数据更好地应用于中药资源。

关键词 大数据 中药资源 数据整合

中图分类号 R28

文献标识码 A

文章编号 1672-397X(2015)04-0067-04

中药资源在生长种植和学科研究中产生了大量的数据,整合挖掘、合理应用这些数据,对于科学管理中药资源有着极其重要的作用,对中药产业的发展和临床用药安全有深远的影响^[1]。维克托·迈尔-舍恩伯格、肯尼斯·库克耶在其编著的《大数据时代》中提出通过搜集、分析海量数据信息,开发针对性应用的新思路。在当前海量数据和信息爆炸的背景下,大数据的思维模式和行动方式已逐渐渗透到各个行业和业务职能领域。将大数据技术应用于中药资源领域,有助于开启中药资源全新的局面。但同时也应警惕过于依赖大数据。大数据自身不可能告知如何从众多数据中提取对中药资源发展有价值的信息,而是依赖于人的判断和决定。如果不能清晰地界定问题,海量数据也可能成为没有价值的信息垃圾,与中药资源发展的追求背道而驰。

1 大数据的内涵

1.1 大数据的定义 目前学界对于大数据的定义,没有统一的定论。根据维克托·迈尔-舍恩伯格、肯尼斯·库克耶的观点,从数据利用价值的角度定义大数据是人们在大规模数据的基础上可以做到的事情,且这些事情在小规模数据的基础上是无法完成的^[2]。大数据的概念,首要特点是“大”,主要体现在数据规模大、数据类型多、价值密度低和处理速度快四个方面。但是大数据不只是数据大,还包括采集数据的工具、平台 and 数据分析系统。大数据的本质是在大量的、复杂的数据集合中获取知识和洞见。其用途可概括为:利用硬件设备、信息技术和计算方法的进步,将不同格式、不同领域的大数据整合成标准统一的

数据源,研究分析挖掘后继而产生巨大的经济社会价值;预测未来某一事件发生的可能性,并据此进行人为干预,使其向着预期的方向发展^[3]。大数据时代带来的变化不仅体现在如何处理海量数据并从中获取有价值的信息,也体现在如何加强大数据技术研发,准确把握未知的趋势,抢占时代发展的前沿。

1.2 大数据对于中药资源数据的意义 中医药秉持“天人合一”的整体性思维,以其整体、宏观、粗放的医学认知实现在宏观整体基础上认知世界、维护人体健康和调整疾病状态,这与大数据的思维具有很高的契合性。中药资源数据主要包括我国中药资源种类、分布、蕴藏量、资源变化趋势、栽培与野生情况、收购量、需要量、质量等^[4]。利用大数据能够创新数据采集方式,改进抽样技术,提高中药资源统计调查的科学性、准确性和实效性。通过数据化整合与共享中药资源数量和质量,可以缩短中药资源调查周期,提高调查效率,节约人力、物力、财力,可以为中药规范化种植、药材提取与深加工等提供技术支撑和保障,并将为科学规划发展中药材产业提供重要参考依据。

1.3 对中药资源数据认识的误区 我国中药资源品种众多,与之相关的生长、种植、流通、利用等环节因素庞杂,由此产生的数据量之多可想而知。事实上,我国当前中药资源信息统计工作面临的问题有:水平总体不高,发展不平衡,统计信息范围较窄,信息仍显粗糙、含混,利用率低等。对数据采集、存储和分析等核心技术关注较少,相比之下更加看重应用领域,甚至出现神化大数据的误区,认为掌握数据和相应的硬件设备就可以分析数据。相比于大数据对各项技术提

基金项目:国家卫生计生委药政司 2014 年度委托研究课题

出的更高、更新的要求,传统的数据采集、收集方式相对套路清楚,简便易学。而现在面对海量数据,如何选择数据,如何判断相关性大的因素,如何建模型等问题尚没有形成完备的技术和方法^[5]。

由于我国尚无采纳数据格式的国家标准,政府职能部门限制使用者获取官方统计数据,导致信息不完整或重复投资。但实际上,有些行业的数据涉及上百个参数,其复杂性不仅体现在数据样本本身,更体现在多源异构、多实体和多空间之间的交互动态性^[5]。以谷歌 2008 年前推出“流感趋势”服务为例,通过分析关键词,准确预测了美国各地流感爆发情况。谷歌综合分析大量实时数据和历史数据,了解不同历史阶段流感爆发情况,以及对对应时期的关键词、频率,判断哪些因素的相关性大,反复试错、调整,才能建立可靠模型。

2 大数据在中药资源数据整合中的应用

2.1 大数据在中药资源数据整合中的应用现状 近年来,运用掌上电脑(PDA),采用遥感(RS)、地理信息系统(GPS)、全球定位系统(GIS)三者集成的“3S”技术手段已逐步显示出良好的应用前景,可以提升和丰富中药资源调研的手段与方法。以正在进行的第四次全国中药资源普查为例,主要通过计算机辅助和专家经验相结合的方式组织实施中药资源普查的数据填报、汇总分析等主要过程。利用计算机实现实施方案能在智能模板填写,通过加入专家的在线审核,可以将信息及时反馈给各级管理部门;计算后台自动实现普查数据完整性、有效性校验以及成果的汇总分析,而通过加入专家设置详细的统计参数,则提高了汇总结果的准确性^[4]。

但由于中药资源的研究过程综合分类学、生态学、地理学、生物化学、天然产物化学和中药学等学科^[6],具有特定的交叉性、集成性和广泛的带动性,决定了这是一项复杂的系统工程,面临着不可规避的瓶颈问题。

中药资源的遥感数据、气象数据、生态系统及地方植物志的植被信息资料、行政区划和土地利用的地理数据等一系列应用系统呈“碎片化”。数据标准不统一,相关知识表达方法不规范,导致这些数据与知识既无法相互兼容,也无法共享,严重制约了我国中药资源研究的深入建设。

当前我国中药材的种植(养殖)多与农业联系在一起,是增加部分地区农民收入的有效途径。随着中药发展的产业化,药农种植(养殖)过程的组织化、集约化,对中药材种植(养殖)的量化决策提出更高的要求,如何进行科学种植(养殖)、精准施药调控、优化品种选择,与更加专业化的决策有密切联系。受成本和技术水平的限制,这些问题都暂时难以解决。

2.2 大数据对中药资源数据整合的机遇 中药资

源是中医药产业发展的源头,其数量和质量一直是制约中医药发展的关键因素。引入大数据的新思维和新思路,不仅有助于促进中药资源研究领域发展,也有利于探索中药资源发展的新方向和新方法。

2.2.1 中药资源的自然生态研究 自然生态环境与中药资源的质量、数量密切相关,是其生态适宜性评价的客观基础。当前我国以“3S”技术为核心的新型监测技术在植物信息提取及植被分类、土壤水分监测、土壤养分监测、病虫害监测等方面具有方便快捷、经济实用的特点^[7]。大数据通过建立一套完善的中药资源地理信息系统,以不同系统采集的海量信息为基础,通过比对识别模型和当前气候条件,进行预测性分析,以用于评估土壤肥力、监测农田边界图管理、预测产量分布、定位虫害情况,并分析控制方法和相关决策,为中药资源的自然生态研究及发展提供有价值的信息。

2.2.2 中药种植规范体系 基于现在多数中药资源明显减少,品质也明显下降的现状,很多地区开始尝试种植中药材以满足日益上涨的市场需要^[8]。然而中药材种植与其他农作物种植相比不仅有产量的要求,也受药效的限制。目前中药材种植缺乏科学化、规范化,药农每年的播种大多是使用自留种子或从其他农户购买种苗,或者到别的产区购买种子种苗,或由自己参加的合作社提供。药材种植管理也较为简单粗放,无法得到专业中药材种植技术指导。大数据可以将中药材种植地的各种信息进行数字化,录入到各类数据库中,自动生成优化的耕作、施肥、灌溉和喷药等指令,结合设计的专家建议系统,进行综合分析,实现科研、科普、引种、栽培、生产等业务工作的信息管理,建立一个权威的网络信息管理系统。

2.2.3 中药材市场流通体系 由于我国尚没有一套权威的中药材行业基础信息,药材收购量、需要量、蕴藏量等基本情况不明确,造成中药材种植户和经营者之间的信息严重不对称,供求信息不畅,市场无法根据供求合理调节价格,让投机分子趁机囤积居奇,炒高中药材价格,制造中药材市场的虚假繁荣^[9]。例如沉香是中医治胃病常用的一味药材,取自瑞香科植物白木香含有树脂的木材,但无序采伐使得野生沉香数量急剧减少,加之资本炒作,使得沉香饮片统货价格从百来元每千克上涨到数千元每千克^[10]。大数据则可以依据生态自然环境因素和政策社会因素,整合预测农产品价格走势、消费需求等相关数据,对及时调整中药数量、质量给予基于数据解读的对策和建议。

2.3 面临的挑战 目前大数据应用还处于探索阶段,它与中药资源现有的创新方法既相互联系,又相互区别。加之其理论框架尚不完备,基础设施、相关制度法律还不成熟,大数据到底能为中药资源研究带来什么变化还令人困惑。

2.3.1 各地中药资源管理主体存在差别 当前各省主管中药资源的部门不同,有的由省卫生厅主管,有的由省中医药管理局管理,也有由省林业厅管理的,而部分省则没有具体管理部门。在有管理部门的省中,分为组织机构职能健全和职能欠缺两类,存在较大的区域差异。在此基础上,要使分散在不同部门的数据得到合理有效应用,明确可用数据、不公开数据都需要政府统筹协调,明确的主管部门履行监管中药资源的职能。

2.3.2 中药资源人才缺乏 我国刚刚引入大数据概念,培养大量掌握大数据挖掘技术和中药资源专业知识的人才不可避免的发展方向。发达国家的部分IT企业已经初步掌握了一些数据分析预测的关键技术,我国相关企业的技术能力还比较落后。另外,虽然中医药高等院校包括综合性大学的药学专业已累计超过300家,每年培育的专业技术人才已逾万人,但是中药资源相关专业在培养人才方面缺乏系统性、整体性,人员流失严重。目前只有23个开设“中药资源与开发”本科专业的院校,一批开设“中药资源”相关专业、方向硕士点、博士点的院校^[1]。

2.3.3 潜在的安全威胁 中药资源不仅包含对其生长环境的种种要求,还包含中药材种植基地选定、品种栽培、采收加工、质量标准等一系列知识信息。虽然现有的保护传统知识信息的法律法规及技术手段能起到一定保护作用,但是大数据的价值并不只是收集海量信息以改变现状,更多用于对收集到的信息进行二次利用,预测和判断未知的行为。因此未来在采集数据前无法穷尽其用途,充分获得信息使用许可,因此对于数据提供者的告知与许可便也失去实质意义,加深对数据隐私、公共机构信息保密的威胁。近年来珍贵药用资源遭到国外医药机构地毯式筛选的事件常有发生,导致我国自主开发中药或天然药物的源泉日益枯竭,同时也威胁了我国生态安全甚至国家安全。在新形势的挑战下,客观冷静面对大数据,使其在合理的规制下发挥作用显得尤为重要。

3 结论及建议

中药资源信息的利用不仅具有生态学意义,更具有国家战略意义。中药资源的开发利用是中药产业的根基,关乎民生和社会稳定,关乎生态环境保护和新兴战略产业发展,是全球竞争中国家优势的体现。将大数据应用于中药资源中意味着很多机遇,但要真正利用好大数据技术,还有许多亟待完善的问题。

3.1 转变思想观念 面对不断升温的大数据热潮,先知、先觉、先行既可能率先受益,也可能首当其冲,关键在于是否能理性分析、冷静观察,找到核心技术、应用模式。大数据要求扎实的基础性工作,这

些工作可能是长期性、前瞻性的,需要逐步推进,很难立竿见影、一蹴而就。因此对当前的大数据潮,既要有积极迎接的准备,也要有警惕过热的思想。通过探索突破传统数据管理的局限,结合中药资源信息管理的特点,在相关法律制度的规范下,促进中药资源学科的发展。

3.2 提高技术水平 有效运用大数据,需要提高计算机和数据管理硬件,增强相应计算系统的功能,培养一批大数据分析运用的服务企业和专业人才。通过建设计算机管理系统,克服中药资源普查项目因经费和人员等原因的非常态化,建立长期固定的资源监控机制,实时把握中药资源的分布及生长状况。

政府可以在构建数据质量分析平台,构架药农、企业和市场间的沟通桥梁方面予以支持。鼓励民间设立专项基金,提高中药行业的技术水平,鼓励将中药资源普查及相关科技成果的转化、应用和推广,为行业提供技术支撑和技术服务^[2]。另一方面,在现有的药用植物园体系的基础上,建设信息共享平台^[3],充分利用数据资源,以收集药典上常用中药、濒危中药为重点,集种质保存、科研、种植培育、科普宣传、旅游五位一体,全面发展。

政府还可以在抽样调查、网络信息技术应用等方面应用大数据的思维和方法,有所突破。空间数据的应用实际上要先于大数据概念的提出,由于空间数据具备大数据的特性,故其本身就是“大数据”^[4]。中药资源的种植(养殖)具有明显的空间特性,是利用空间数据的主要领域。尝试利用航拍技术获取具有空间特征的中药材种植(养殖)区域、居住区域等对象的高清影像,结合已知的普查信息,构建空间抽样框。并通过“3S”技术对样本实行遥感和观测,使调查对象实现可视、可测、可追溯。在网络信息技术方面,可以逐步对中药材价格调查进行现场数据采集,通过基层统计报表传送终端设备,基层统计人员手持设备到现场,直接将采集的信息上传。

3.3 完善法律制度 大数据为监测中药资源提供了便利,但同时对保护中药资源相关传统知识的手段提出新的挑战。原有的数据隐私保护模式不能再起到良好的作用,因此需要完善数据隐私保护,建立健全公共机构信息公开等立法修法工作。我国于2012年出台的《关于加强网络信息保护的决定》、《信息安全技术公共及商用服务信息系统个人信息保护指南》都对大数据时代的个人信息信息管理进行了原则性的规定^[5]。在未来的数据隐私保护规制中,应进一步明确数据用途,包括数据可以进行商业化应用的范围、应用范围界定标准、数据滥用承担的责任等。对于一些危险性较大的项目,应设立专门规章,规范数据使用者评估风险,降低潜在伤害。

(下转第72页)

调经理气为主,用桃仁丸(桃仁、大黄、朴硝各 3g,蛭虫 5g,先将三味同熬好后入朴硝化为丸,如桐子大,每服 5~7 丸,五更时白汤下)。郑氏区别了癥瘕之不同,一般多先有气聚成瘕,日久则血聚为癥,但两者都与气血密切相关。大致可分为瘀热内阻、肝郁血瘀、阴虚痰瘀、寒瘀互阻、肾虚肝郁五个类型,其治疗方法也灵活多变,有清热化瘀、疏肝理气化瘀、滋阴降火化瘀、温经散寒化瘀、补肾解郁化瘀等治法。在用药上也强调瘀热湿浊的病理因素,故常在方中配入炙甲片、鳖甲、淡昆布、海藻、牡蛎、王不留行等咸寒辛润软坚散结之品。在化瘀的同时,郑氏也注重气血的调节,气血是脏腑的物质基础,气血逆乱则易导致脏腑功能的虚衰,月经的失调。在治疗癥瘕的过程中郑氏并非概以行气活血而治之,常根据临床辨证分别予理气、补气、养血、调经之法以获良效。

7 育肾保精,温通奇经治不孕

郑氏提出女子不孕主要是由于肾气阴精不足,或冲、任、督三脉气血失调所致。《女科要旨》云:“妇人无子,皆有经水不调,经水所以不调者,皆由内有七情之伤,外有六淫之感,或气血偏盛,阴阳两乘所致。”肾精是女性怀孕的物质基础,肾气旺盛,精血充沛,任脉通畅,冲脉盛实,督脉温通,两精相搏,才能受孕。在治疗不孕过程中,郑氏注重调节奇经的重要功能。奇经主要功能一是沟通十二经脉之间联系,对十二经脉气血有蓄积渗透作用。奇经中尤以任、冲、督三脉对生殖影响较大。郑氏认为任脉总任一身之阴经,起于胞中,濡养女子胞宫,使胞宫阴精充盛。冲脉能调节十二经气血使女子子宫经血按时

蓄溢,月经按时来潮。督脉总督一身之阳气,行于背脊,上行入脑,与脑、脊髓、肾有密切联系。这些奇经的血气流畅与月经的来潮以及月经周期中的排卵都有密切关系,临床施方中常配伍龟版、鳖甲、鹿角霜、川断等药。同时也立体动态地观察到女性不同生理期的子宫气血的内在变化。根据女性月经期、排卵期、卵泡期、黄体期不同生理期,肾阴、肾阳的相互转化,选用不同中药调整子宫内环境,激发体内生殖功能,以促进怀孕。故在各个生理期分别用不同的药方,通补兼施,调整月经的正常运行,正所谓“种子之法,皆在调经之中”。

8 结语

郑氏妇科代代相传至今,辨证思路广泛,组方灵活,用药轻简。著作文风活泼,问答歌诀,治病保健,无所不及。不但为传承者提供了宝贵的临床经验,而且也为当今各种妇科疑难杂症的治疗拓宽了思路,为今后的研究与发展奠定了坚实的基础。

参考文献

- [1] 马一平. 昆山郑氏妇科二十九代世医考. 中华医史杂志, 2000;2
- [2] 林士毅, 周坚, 刘时觉. 郑氏女科与《坤元是保》. 浙江中医杂志, 2013, 48(2): 94
- [3] 任宏丽, 刘玉良. 《女科万金方》中四物汤的灵活运用. 山西中医, 2008, 24(8): 40

第一作者: 陆爱芳(1978-), 女, 本科学历, 主治医师, 从事中医妇科疾病临床诊疗工作。
laaifang@sina.com

收稿日期: 2014-11-20

编辑: 傅如海

(上接第 69 页)

参考文献

- [1] 潘文, 程涛, 牛崇信, 等. 大数据时代中医药信息的应用. 中国中医药图书情报杂志, 2014, 38(1): 2
- [2] 维克托·迈尔-舍恩伯格, 肯尼斯·库克耶. 大数据时代. 杭州: 浙江人民出版社, 2013: 228
- [3] 蔡书凯. 大数据与农业: 现实挑战与对策. 电子商务, 2014(1): 3
- [4] 郭兰萍, 陆建伟, 张小波, 等. 全国中药资源普查技术规范制定. 中国中药杂志, 2013, 38(7): 937
- [5] 郭贺铨. 大数据时代的机遇与挑战. 求是, 2013(4): 47
- [6] 周荣汉. 中药资源学. 北京: 中国医药科技出版社, 1993: 3
- [7] 郭兰萍, 黄璐琦, 蒋有绪. “3S”技术在中药资源可持续利用中的应用. 中国中药杂志, 2005, 30(18): 1397
- [8] 赵文吉, 李敏, 黄博. 中药材种子种苗市场现状及对策探讨. 中国现代中药, 2012, 14(3): 5
- [9] 于志斌. 中药材市场调研分析报告. 中国现代中药, 2012, 14(1): 51
- [10] 李瑶. 中药材源头之忧. 医药经济报, 2014-7-28(5)

- [11] 段金廛, 周荣汉, 宿树兰, 等. 我国中药资源科学发展现状及展望. 自然资源学报, 2009, 24(3): 378
 - [12] 陈士林, 肖培根. 中药资源可持续利用导论. 北京: 中国医药科技出版社, 2006: 56
 - [13] 李标, 魏建和, 王文全, 等. 推进国家药用植物园体系建设的思考. 中国现代中药, 2013, 15(9): 721
 - [14] 马建堂. 大数据在政府统计中的探索与应用. 北京: 中国统计出版社, 2013: 147
 - [15] 马晓亭. 大数据时代基于服务等级协议的图书馆读者隐私感知与保护研究. 情报理论与实践, 2014, 37(4): 57
- 第一作者: 田虹(1990-), 女, 硕士, 研究方向: 药事管理、卫生法。

通讯作者: 田侃, 教授, 博士生导师. tiankane@qq.com

收稿日期: 2014-10-09

编辑: 吴宁