

DOI: 10.13288/j.11-2166/r.2016.20.021

百 家 园

基于大数据与中医药研究现状对慢病防治的启示

濮文渊¹, 凌云^{1*}, 周春祥^{1,2}

(1. 南京中医药大学基础医学院, 江苏省南京市栖霞区仙林大道 138 号, 210046; 2. 南京明基医院)

摘要 随着人类生产方式、生活方式、行为方式的改变, 慢病已成为目前危害人类健康的重要公共卫生问题, 传统中医药也应结合现代先进科学技术, 发挥特色优势。就当下新兴的大数据技术在中医药领域应用现状进行综述分析, 提出推进中医药大数据化过程中可能存在的问题及相应对策, 为现代慢病防治工作提供借鉴与参考。

关键词 慢病; 大数据; 数据挖掘; 数据仓库; 中医现代化

随着社会产业化及高科技、商品经济的发展, 人类的生产方式、生活方式、行为方式均已发生改变, 社会竞争剧烈、工作节奏加快, 人类的疾病谱也随之改变。早在 20 世纪末, 有学者统计分析得出我国自 1957 年以来城市人口的主要死因已发生了根本改变。从疾病谱的构成来看, 解放前占总死亡率一半以上的肺结核、急性传染病、消化系统疾病等已不再列入 1993 年我国城市人口前五位主要死因, 取而代之的为恶性肿瘤、脑血管病、心脏病^[1]。

1 现代疾病谱变迁, 慢病时代到来

慢病主要是指以心脑血管疾病、糖尿病、恶性肿瘤、慢性呼吸系统疾病等为代表的慢性非传染性疾病的总称^[2]。据 2011 年世界卫生组织 (WHO) 调查结果显示, 2008 年因慢病死亡的人数高达 3600 万, 而在中国等发展中国家慢病发生率更高, 约占世界慢病总死亡数的 80%, 造成的疾病负担占国家卫生总费用的 70%^[3], 是目前危害人类健康的重要公共卫生问题。积极控制危险因素和预防疾病成为降低慢病发生发展的根本措施。然而目前我国慢病防治工作主要依赖于流行病学研究, 群体干预技术及优化方案等传统公共卫生手段, 缺乏根据因人而异的“个体化”干预措施, 导致我国慢病防治工作仍难以取得突破性进展^[4]。

2 大数据为中医药现代化带来机遇

大数据作为信息技术产业高速发展的产物, 已逐渐渗透到人们工作和生活的各个领域。相比于传统的数据库应用, 大数据具有数据量大、查询分析复杂等特点。它可以将规模巨大、无法通过目前主流软件工具的资料量, 在合理时间内摄取、管理、处理, 并整理成高价值资讯, 为企业经营者做出前瞻性决策服务^[5]。大数据对数据和信息的分析提升到了前所未有的高度, 因此, 如何充分认识、挖掘和利用中医药大数据, 从中发现有价值的信息, 使之为临床实践和决策服务, 是中医药现代化的突破点^[6], 大数据技术已在中医药研究的多个分支领域内获得初步应用。

2.1 中医古籍研究

陈肇文^[7]在中医古籍治疗中风的医案中, 运用频次分析和关联分析的方法, 得出古代名医治疗中风最常用的 7 味中药、9 个药对。宋咏梅等^[8]对历代治疗不寐方剂中的中药进行频次分析得知, 酸枣仁是所有方剂中使用频率最高的药物, 使用频率较高的有人参、茯神、麦冬、半夏、远志、熟地黄、生地黄、柏子仁等。周德生^[9]集成了明清时期记载“津液亏损”的病案 573 例, 采用 R 型系统聚类分析方法, 确定了每个亚型的辨证用药规律。张启明等^[10]在 Visual Foxpro 平台上, 通过对历代中医医案文本数据的规范化、结构化处理, 收录了历代医案 22 459 首, 数据量达到 2 亿。

基金项目: 国家自然科学基金 (81273639)

* 通讯作者: yunlizy@163.com

2.2 针灸疾病谱研究

杜元灏等^[1]统计并分析 20 世纪具有代表性的五部针灸临床著作和 WHO《世界卫生》刊登的有关针灸文献,初步得出针灸临床病谱包括 414 个病症。黄琴峰^[1]回顾 1954 年至 2005 年有关消化系统针灸文献 5219 篇,从多角度数据分析并结合临床应用,推荐针灸治疗的消化系统有效病症为呃逆,常用有效病症为腹痛、腹泻、便秘、呕吐等。齐丽珍等^[2]对新中国成立以来 76 种中医期刊的 56 267 条信息进行整理、归纳和分析得出,针灸临床所涉及的病症数达 972 种,其中以神经系统、五官科、运动系统、消化系统、皮肤科的病症数较多。庞勇等^[3]回顾在广西中医学院第一附属医院针灸科住院治疗的 1388 例患者的病种与疗效、治疗方法、病源等关系,得出神经、精神、循环及运动系统疾病是针灸治疗的主要病种,而其中脑病相关疾病所占比例最高。

2.3 临床病种常见证候研究

余海滨等^[4]通过构建临床数据采集、数据仓库、数据挖掘三大平台开展艾滋病、慢性阻塞性肺疾病(COPD)证候分布研究,结果显示,艾滋病并发胃脘痛患者以脾胃虚寒证最为多见,其次为肝气犯胃证、饮食伤胃证; COPD 急性加重期证候分布以痰热壅肺证、肺肾气虚证、痰瘀阻肺证为主要证候。马金辉等^[5]就全国 16 家三级医院中 30 034 例高血压病住院患者档案开展数据分析,得出高血压病患者多合并冠心病、血脂蛋白紊乱血症等,其中中医证候以痰瘀互结证和肝肾阴虚证居多。李贵华等^[6]选择全国 17 家三级医院中 84 697 例冠心病住院患者进行数据分析,得出冠心病患者最常见合并的疾病为高血压病、糖尿病、脑梗死、血脂蛋白紊乱血症等;其中中医证候以气阴两虚、气虚血瘀最为多见,血瘀是比例最高的证候要素,占 79.97%。

2.4 临床慢病常用中药研究

胡金亮等^[7]收集 2011 年 8 月至 2013 年 3 月河南中医学院第一附属医院 5000 份电子病历,运用水晶报表进行数据挖掘分析,得出痰热壅肺证、痰瘀阻肺证、肺肾气虚证的处方分别对应的中药均有甘草、陈皮、地龙、紫苏子、黄芩、红花等。余海滨等^[4]采集名老中医治疗艾滋病并发咳嗽的处方 523 份,使用中医临床复杂网络分析系统进行用药规律研究,得出其核心药物有甘草、款冬花、黄芩、半夏、黄芪、黄连、防风、麻黄、冬瓜仁等。

3 大数据应用于中医药领域研究现状分析

通过前文文献综述,可以看到已有不少学者在中医古籍、针灸、临床常见病症、常用中药等方面运用大数据技术做了有益探索,大数据为中医药领域带来了重大变化。中医药大数据化的优势在于:1)就中医古籍而言,大数据时代打破了古今年代、地域界限和流派之争,使得面向主题的全面信息的存储和检索变得简单和可能。通过收集古籍中大量的、复杂的、模糊的、不确定的个体经验,分析数据集合从而获得有用的信息和知识的过程^[8-10],在数据挖掘后呈现总体的规律性。2)就针灸而言,大数据技术将现代针灸疾病谱研究带入了新的发展阶段,有学者根据临床针灸的实际情况,将针灸临床疾病谱划分为 4 级,使得针灸与时俱进,规范了现代针灸的应用范围,更有利于针灸的推广与传播^[11]。3)在临床慢病防治方面,通过大数据分析大量临床患者的病历资源,总结慢病的常见证候及常用方药,不仅有利于慢病防治经验累积,也有利于及早对患者进行干预,降低疾病风险。

但遗憾的是,目前大数据与中医药现代化的结合尚处于初步探索阶段,仍存在很多困难和问题亟需解决。如全国近 800 家三级医院,绝大部分医院的数据资源尚未实现共享;并且目前大部分临床研究都较为单一,仅局限于对某种疾病大数量病例的统计与分析;在数据挖掘过程中也缺乏统一规范,不利于后期对现有研究成果进一步统计总结得出规律性结论。作为中医特色的个性化诊疗及宝贵的名老专家用药经验,也可能随着数据规范化而流失。

为此我们建议,在应用大数据技术促进中医现代化发展的过程中,可从以下 7 个方面做些工作:

- 1) 相关政府机构在国家层面上建立数据中心,尽快完善各等级医院数据库与数据库之间的整合;
- 2) 研究制定中医药临床数据共享管理办法,鼓励各医疗单位共享现有平台资源,打破各自为阵的传统观念;
- 3) 统一中医临床实践、古籍文献等方面的数据录入规范,统一数据采集办法,为后期整合数据资源、挖掘更深层次的信息打下基础;
- 4) 培养既懂中医又熟悉计算机大数据技术、掌握统计技能的复合型人才,与高科技行业建立紧密合作关系,保持中医药理论本色,保障大数据先进技术;
- 5) 建立系统完整的疾病证候分类模型,制定行业公认的符合现代临床实际的“证”的标准,界定从复杂症状中去粗取精提取出关键症状的判断方

法, 确立临床常见多证兼见情况的处理原则; 6) 适当规范古籍中一词多义、多词同义现象, 完善古籍信息数据库, 避免信息误读漏读, 提高提取信息价值; 7) 利用已成熟的云计算技术和计算平台, 设计集诊疗与健康管理为一体的慢病数据管理平台, 通过分析患者平时和治疗期间的生理病理数据, 既可帮助医生在临床中提高药物疗效, 又可对患者实时监测身体状况, 督促患者调整不良生活习惯, 降低疾病发生和复发风险。

4 结语

东汉末年, 张仲景通过研求众方并结合自己的临床实践, 完善和丰富了经方类证系统并使理、法、方、药一线贯通, 总结出当时“伤寒十居其七”的疾病谱, 最终著成《伤寒杂病论》。清代乾隆以后, 叶天士结合当时临床常见的呼吸系统传染病、猩红热及斑疹伤寒等传染性疾谱, 提出“温邪上受, 首先犯肺, 逆传心包”, 以清灵通通为治疗原则, 用芳香祛秽、甘寒生津、育阴清热、清营凉血等治法, 为温病危重急症的治疗独辟蹊径, 从而著成温病学说奠基之作——《温热论》。其实历代经典医家著作都是基于当代临床“大数据”凝练而成的精华。

而今大数据时代到来, 我们可以充分利用高科技的力量, 更新现代疾病谱, 并根据时代、环境的改变做出相应的调整。我们不仅可以利用现代科技分析现代人患病的常见证候, 对比古今证候异同, 丰富现有证新的症状, 总结出新时代的疾谱, 还可以从宏观视野分析当下慢病时代的复杂病因, 甚至找出暗藏在各种慢病背后潜在的共同病因及不同慢病之间的相关性。在此基础上, 我们既可以根据当代疾谱变化和患者证候分布来指导医者临床用药处方思路, 实现“从临床中来, 到临床中去”的理论回归; 又可以针对性地指导现代人防病养生, 提供个性化药膳和养生要点, 在生活中贯穿“未病先防, 既病防变”的中医核心观念; 甚至可以根据临床常见症、证的用药规律, 指导中草药的种植与生产, 大大节约药材资源, 减少不必要的资源浪费。

参考文献

[1] 田玮, 王雅珍. 建立适应 21 世纪医学教育模式的思考 [J]. 医学与哲学, 1998, 20(1): 55-56.

- [2] 贺婷, 刘星, 李莹, 等. 大数据分析在慢病管理中的应用研究进展 [J]. 中国公共卫生, 2016, 32(7): 981-984.
- [3] HUNTER DJ, REDDY KS. Noncommunicable diseases [J]. N Eng J Med, 2013, 369(14): 1336-1343.
- [4] 潘文, 程涛, 牛崇信, 等. 大数据时代中医药信息的应用 [J]. 中国中医药图书情报杂志, 2014, 38(1): 2-4.
- [5] 刘保延. 大数据绘制当代中医航海图 [N]. 中国中医药报, 2013-06-05(3).
- [6] 陈肇文. 数据挖掘技术在古代名中医中风医案之应用研究 [J]. 中华中医药学刊, 2008, 26(10): 2254-2257.
- [7] 宋咏梅, 崔利锐, 李军伟. 古代失眠方药的文献分析 [J]. 山东中医药大学学报, 2012, 36(5): 392-394.
- [8] 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, 14(4): 12-13.
- [9] 张启明, 王永炎, 张志斌, 等. 中医历代医案数据库的建立与统计方法 [J]. 山东中医药大学学报, 2005, 29(4): 298-299.
- [10] 杜元灏, 肖延龄. 现代针灸临床病谱的初步探讨 [J]. 中国针灸, 2002, 22(5): 59-62.
- [11] 黄琴峰. 消化系统针灸疾病谱现代文献计量分析与评价 [J]. 中国针灸, 2007, 27(10): 769-772.
- [12] 齐丽珍, 黄琴峰, 黄颖. 从现代中医期刊透视针灸疾病谱 [J]. 上海针灸杂志, 2006, 25(11): 46-50.
- [13] 庞勇, 李保良, 农泽宁. 针灸病房患者 1388 人次临床资料分析 [J]. 中国针灸, 2002, 22(5): 55-57.
- [14] 余海滨, 符宇, 李卓, 等. 基于临床科研信息共享系统开展中医临床研究的探索 [J]. 中医杂志, 2013, 54(24): 2092-2094, 2109.
- [15] 马金辉, 王志飞, 谢雁鸣, 等. 真实世界大数据 30034 例高血压病住院患者中西医诊疗规律初探 [J]. 中国中药杂志, 2014, 39(18): 3435-3441.
- [16] 李贵华, 姜红岩, 谢雁鸣, 等. 基于大数据 84697 例冠心病中医证候及其中西药使用分析 [J]. 中国中药杂志, 2014, 39(18): 3462-3468.
- [17] 胡金亮, 张润顺, 周雪忠, 等. 基于临床医疗科研信息共享系统中医病-证-中药关联分析方法的研究 [J]. 世界科学技术-中医药现代化, 2013, 15(5): 905-909.
- [18] OLSON DL, DELEN D. Advanced data mining techniques [M]. New York: Springer Publishing Company, 2008: 5-7.
- [19] MEHMED KANTARDZIC. Data Mining: Concepts, Models, Methods, and Algorithms [M]. Hoboken: A JOHN WILEY&SONS, INC, 2011: 1-9.

(收稿日期: 2016-02-27; 修回日期: 2016-03-16)

编辑: 黄 健