

DOI: 10.13288/j.11-2166/r.2018.13.007

中医医案分析方法评述

何彦澄, 肖永华*, 闫璞, 陈净殊, 王世东

(北京中医药大学东直门医院, 北京市东城区海运仓 5 号, 100700)

摘要 中医医案是最能体现医生学术思想与辨证论治的记录。通过对近年来中医医案分析类文献的回顾, 基于中医医案具有显性知识与隐性知识的知识特点, 以及医案全数据性、难以规范化的数据特点, 从群案及个案分析研究现状出发, 结合中医个体化辨证论治优势, 提出个案数据分析的必要性及可行性。

关键词 中医医案; 群案分析; 个案分析; 隐性知识; 中医智能处方分析系统

中医医案是医生临床诊治患者的记录, 最能体现医生的学术思想和辨证论治艺术技巧^[1]。因此, 从中医继承和学习的角度而言, 在浩瀚的中医典籍海洋中选取医案进行研究, 不失为是一条行之有效的途径。进行中医医案研究, 实质也是针对中医医案的分析方法的探索, 适当的分析方法能够大幅度提高中医继承学习的有效性。现就目前中医医案分析方法现状简要评述如下。

1 中医医案知识特点

根据波兰尼对显性知识与隐性知识的定义^[2], 中医知识可以明确分为显性知识与隐性知识两部分。隐性知识的存在是中医神韵“可以意会, 难于言传”^[3]的重要原因, 即所谓“医者意也, 意者医也”, “中国医学理论的神秘性、治疗方法的灵活性、医家的悟性, 都只能以一个‘意’字来体现”^[4]。隐性知识是构成中医药传统知识的重要因素, 是中医传承的命脉、发展的基本动力和中医知识不断创新的源泉^[5]。

中医医案作为中医理、法、方、药综合运用的具体反映形式^[6], 其中也包涵显性知识与隐性知识。医案中记录的患者症状、体征和医生处方是显性知识; 医案中蕴含的医家学术思想、临床经验, 甚至是其医德和治学, 均为未编码的知识, 属隐性知识。是否能更为完善地继承和挖掘隐性知识, 是中医药继承工作成败的关键所在。许多老中医药专

家学术思想和经验因其隐性知识的特性与传统中医传承方式的限制而得不到传承, 导致一些特色诊疗技术和方法濒临失传, 所以将中医医案中的隐性知识系统清楚地表达出来, 使之实现编码化, 形成相对客观、容易学习的显性知识, 中医学才能更好地传承。

2 中医医案研究的数据特点及分类方法

2.1 医案数据特点

2.1.1 全数据性 中医信息具有全数据性的特点, 即在观察某一对象时, 尽可能全地收集与该事物相关的信息, 包括本质的、现象的, 稳定的、不稳定的。这些数据达不到海量的程度, 但对某一具体事物来说是全面的, 能够从中发现事物相关关系的细节^[7]。因此, 在对医案使用数据挖掘手段时, 数据全面性较数据量更为重要。

2.1.2 难以规范化 统计归纳与数据挖掘方法要求一定的数据规范性, 而中医医案的描述方式具有一定的主观性、随意性, 存在缺损数据多、噪声数据多、信息冗余、数据动态变化等缺点^[8], 尤其是古代医案缺乏规范与条理^[9], 导致中医医案的数据规范化一直是医案数字化研究中的难点问题^[10]。仅从数理角度挖掘出的医案信息缺少对医案中隐含知识的发现和解析^[11], 不能直接指导临床与科研。

2.2 医案研究常见分类方法

2.2.1 个案与类案 中医医案从数量上分为个案和类案。个案是指一两个医案的记载, 以其数量少、个性化为特征; 类案又称为群案, 是指一系列的同类的医案, 多按照某一位医家、某一类病症

基金项目: 国家中医药管理局全国名老中医药专家传承工作室建设项目; 国家中医药管理局委托办事经费; 北京中医药薪火传承“3+3”工程建设项目(JB059)

* 通讯作者: water_aqua@126.com

或治法归类, 个案是群案的基础和典型代表。

2.2.2 验案与误案 从临床疗效上可分为验案与误案。验案指治疗效果为有效、显效或痊愈的医案; 误案指因判断失误, 使得治疗效果为无效, 甚至使病情进一步恶化的医案。

2.2.3 古代医案与近现代医案 根据时间分为古代医案与近现代医案。古代医案自淳于意^[25] 则诊籍始, 至清代为其鼎盛时期, 以中医临床证治、中医经典理论分析为主, 医案体例基本相同, 行文缺乏规范, 研究方法有限, 具有朴素性与单一性的特点; 近现代医案自西方科学和医学知识广泛传入后至今, 在中医学基础上加入西医学内容, 现代则采用西医病案体例, 加入中医辨证论治内容, 变得更加细致与规范^[11]。

根据不同的特点, 医案的分类方式有所不同。近十年医案研究状况显示, 医案研究就研究对象分类而言可分为以医类案、以病证类案、以方类案三种^[12], 其研究对象均为类案, 即群案, 结合医案分类方式, 本文从群案及个案角度分析中医医案研究情况。

3 群案及个案研究现状

在中国知网以“医案”为主题词检索, 共得到 17447 条结果, 文献发表年限最早为 1954 年, 其中 16951 条结果为个案报道、综述等相关文献, 不在本文探讨范围内; 以“医案研究”“医案分析”为主题词检索得到 496 条结果, 仅占医案类文献的 2.93%; 以“医案”为主题词, “数据挖掘”“统计”为关键词检索得到 483 条结果, 最早发表年限为 2002 年, 其中 431 条为群案研究, 占医案研究的 86.8%, 1 条为个案研究, 仅占医案研究的 0.2%, 50 条为综述和数据挖掘技术相关研究。由目前得到的数据可知, 现代医案研究以群案研究为主。

3.1 群案研究

3.1.1 群案常用分析方法 目前针对群案的研究主要有统计归纳和数据挖掘两种类型, 统计是收集、处理、分析、解释数据并从数据中得出结论^[13], 归纳是指从许多个别的事物中概括出一般性概念、原则或结论的思维方法。数据挖掘是一种多学科综合的产物, 综合利用人工智能、机器学习、模式识别、统计学、数据库、可视化技术等, 自动分析数据并从中得到潜在隐含的知识, 从而帮助决策者做出合理并正确的决策^[14]。数据挖掘注重知识发现的整个过程, 包括数据如何存储与访

问, 算法如何扩展以有效适合大数据量, 结果如何被解释与可视化, 如何支持人机交互; 统计关注的是统计模型有效性的数学理论依据, 而数理统计理论在数据挖掘各任务中有广泛应用^[15]。统计分析属于数据挖掘方法的一部分^[16], 数据挖掘是比统计归纳更高级、更复杂的一种研究方法^[17], 目前应用有频数分析、相关分析、关联规则、聚类分析、主成分分析与因子分析、Logistic 回归分析、粗糙集理论、人工神经网络、贝叶斯网络、典型相关分析等^[8,18]方法。统计归纳和数据挖掘对医案的规范化要求较高, 主要用于对医家临证的共性规律的挖掘^[17], 是适用于群案的分析研究方法^[19]。目前使用数据挖掘方法, 对医案内部信息进行规律性发掘, 取得了一系列如证候规律、方剂规律、用药规律等成果^[9,20-23]。

目前数据挖掘技术在中医领域还处于起步阶段^[24], 上文提到的数据挖掘手段, 频数统计分析、相关关系分析等统计方法较为常用, 通过定量分析得出定性结论, 理论上比较客观和具有说服力。但实际上, 统计方法的应用通常会依赖一个潜在概率模型。例如研究“证-症”关系中引用的 KNN 算法的工作前提是每个症状均对证候具有相同的贡献度^[25], 而这样的逻辑基础对于时有“但见一症便是”的灵活的中医辨证而言, 显然难以适用。贝叶斯方法与 KNN 算法比较, 相对解决了易受无关症状干扰的问题, 对于古代方剂的证-症判定有一定效果^[26], 但中医临床医案与方剂相比, 其包含的证候与证型可能更加复杂。聚类分析、人工神经网络的核心目的是希望借助计算机技术中的文本挖掘方法, 更深层次地挖掘隐藏在大量医案后面的规律, 但医案文本中常包含大量医家的主观认识, 挖掘结果并不能避免主观认识的介入。

3.1.2 群案分析常见问题 分析结果的不必要性: 对每一个数据都进行分析而不进行筛选, 导致分析的数据里存在一些无意义或者无特异性的数据, 分析结果也就无意义或无特异性了。如黄晓红^[28] 研究历代医家治疗足三阴虚损组方思想, 其统计分析结果之一为三阴虚损临床以内科与外科病多见, 但其临床疾病只分内科疾病与外科疾病, 所以这个结果不具有临床意义。徐玲艳^[29] 对沈炎南药对配伍数据挖掘结果之一是在药对的强关联规则中, 与甘草配对出现的支持度高, 但甘草为调和诸药之品, 自张仲景始, 历代医家喜用甘草者张袂成阴, 用此结果反映医家学术思想不具特异性。

与其研究对象的主观认识有偏差: 直接对数据

结果进行分析,得到的结果与其研究医家的主观思想有偏离。如刘蕊^[6]分析许润三教授治疗女性排卵障碍性疾病的数据挖掘结果显示,许润三对排卵障碍性疾病辨证以肾虚肝郁证为主,治疗以补虚与活血化瘀为主,而许润三所著相关文献^[1]却提到治疗输卵管阻塞主要以疏肝理气、活血化瘀、润管通管为主。郑志鹏^[2]研究干祖望教授诊治慢性咽炎的用药规律,得出干祖望治疗慢性咽炎总以中土脾胃为中心,治法以培土生金、培土补脾为主,用药多以补益脾胃、补肺生津为主,干祖望指导的相关文献^[3]中则提到慢性咽炎主要是津液不足、虚火上炎所致,与肺、肾、脾脏腑功能虚弱关系密切。这样的偏差不一定是没有意义的,但是在对数据结果进行分析时,是否也应该与所研究医家的主观思想相结合?

重视普遍规律而忽略个体:统计归纳与数据挖掘通常用于分析共性规律,因此,在得到数据结果时多分析占比多的部分,而少论述占比少的部分。如张慧^[4]研究清代情志相关病证医案时因记载单一脉象的有 27 条,在 14 种单一脉象中以弦脉出现频次最高,由此得出气机不利、肝失疏泄是情志相关性病证的主要病机的结论,但是在其单一脉象分布频数表中还有数脉、涩脉、细脉、沉脉等脉象,它们所体现的病机在情志相关病证中是否也值得注意?凌雅睿^[5]采用数据挖掘方法总结张书信教授治疗慢性便秘的经验为慢性便秘主要病机以本虚为主,但在其数据中,还存在 25% 以气机郁滞、胃肠湿热、肠胃积热和寒凝气滞等实性病机导致的慢性便秘,而主要结论中没有体现这部分内容。

3.2 个案研究

个案更能体现中医的辨证论治思维^[1],对个案的研究方法主要是领悟式分析方法^[6],而非更具说服力的数据挖掘方法。个案分析在医案文献中占大约 75%^[12],目前其分析方法还停留在传统的研究方式,效率不高,主观性太强,且没有客观数据支撑,无法取信于人。使用数据分析研究个案仅占现代医案研究的 0.2%,其中尚有大片空白。

3.2.1 个案数据分析的必要性 个体虽然存在差异性,但是其生理的相似性存在,个体中包含着一一般的规律,深入研究个案能起到以近知远、以今知古、以所见知所不见的重要作用^[7]。中医医案具有四大特点,即突出创新、实用性、重视个案及较强文学性^[8]。不同于西医的注重群案分析^[1],认为个案是群案的缩影^[9],由于中医临床注重辨证论治、三因制宜,各个医家、流派的学术传承不

同,使得中医医案往往是一人一方,突显出个体化的特点,尤其是一些名家的个案,往往具有较强的代表性、较高的特殊性,能反映其独特的临证经验和学术思想^[1],加之东方传统思维方式的特点是重视事物的不确定性,故中医更注重个案的价值^[1]。但随着西医理念的渗入,中医备受推崇的个案越来越受忽视。针对个案分析,可弥补群案数据挖掘热潮中对个性的忽视;通过对个案的研究,也可以推导普遍规律,个性与共性能得到互相印证,因此,个案的数据挖掘具有必要性。

3.2.2 中医智能处方分析系统 中医医案的处方是研究医生经验最客观的数据^[8],而以处方作为切入点、以个案为研究对象的中医智能处方分析系统即可为个案数据挖掘提供技术支持。如黄苗等^[9]对 1 例慢性肾炎患者的量化研究所得到的系统分析结果与吕仁和教授的学术思想大致符合。由此可见,个案的数据分析是具有可实现性的。

4 小结

中医知识以隐性知识为要,中医医案有着重视个案的特点。对医案的数据挖掘分析,人机结合,是实现中医隐性知识显性化的手段之一。对中医医案的数据挖掘研究因西医思想的影响,目前更多见对群案的数据挖掘,而群案数据分析存在数据规范化程度不够、分析结果准确性不足、挖掘技术有待改善等问题。对个案的研究多是领悟式分析,缺乏证据证明,个案数据分析目前仍少有涉及。那么是否个案数据分析能在一定层面上弥补群案数据分析的缺点呢?目前还没有更多的证据能支持这一点,我们提出这样一个问题,留待后续探讨。

参考文献

- [1] 孟庆云. 宣明往范,昭示来学:论中医医案的价值、特点和研究方法 [J]. 中医杂志,2006,47(8):568-570.
- [2] 夏登杰. 从缄默知识角度谈中医教育改革 [J]. 中医教育,2008,27(5):13-17.
- [3] 汤莉娜,申俊龙,赵宁. 有效提高中医隐性知识传承的策略研究 [J]. 中医杂志,2014,55(4):451-454.
- [4] 刘洪. 基于隐性知识的中医知识保护探析 [J]. 南京中医药大学学报(社会科学版),2013,14(4):180-182.
- [5] 范宇鹏,毛炜,吕玉波,等. 从隐性知识管理角度探讨名医工作室在中医传承工作中的作用 [J]. 中医药管理杂志,2009,17(3):193-195.
- [6] 王佑华,陆金根,柳涛,等. 中医医案中的知识发现研究 [J]. 中西医结合学报,2007,5(4):368-372.
- [7] 崔蒙,李海燕,雷蕾,等. “大数据”时代与中医药“知识密集型”数据 [J]. 中国中医药图书情报杂志,2013,37

- (3): 1-3.
- [1] 于凌. 古代医案数据库研究方法探讨 [J]. 江苏中医药, 2009, 41(8): 4-6.
- [2] 于琦, 崔蒙, 刘丽红, 等. 中医医案数据库研究现状 [J]. 中国数字医学, 2013, 8(3): 71-74.
- [3] 于琦, 王蕊, 李敬华, 等. 中医医案数据的规范方法探索 [J]. 中国数字医学, 2014, 9(10): 73-75, 81.
- [4] 徐世杰, 王国为. 中医医案的特点及其价值 [J]. 北京中医药大学学报(中医临床版), 2013, 20(5): 7-11.
- [5] 刘贤亮, 黎创, 毛炜. 近十年中医医案研究进展 [J]. 中华中医药杂志, 2012, 27(8): 2132-2134.
- [6] 贾俊平. 统计学 [M]. 北京: 中国人民大学出版社, 2009: 2.
- [7] 王梦雪. 数据挖掘综述 [J]. 软件导刊, 2013(10): 135-137.
- [8] 周忠眉. 数据挖掘与统计理论 [J]. 漳州师范学院学报(自然科学版), 2006(1): 23-26.
- [9] 董芸. 数据挖掘与统计分析 [J]. 统计与决策, 2010(9): 2.
- [10] 古求知, 罗琼, 柳长华. 中医医案类文献的分析挖掘研究 [J]. 辽宁中医杂志, 2011, 38(10): 1976-1977.
- [11] 章怡祎, 陆金根, 顾仁樾, 等. 统计分析方法在中医医案研究中的应用 [J]. 中国中医药信息杂志, 2008, 15(3): 8-9.
- [12] 刘浩. 中医医案的价值与特点 [J]. 实用中医内科杂志, 2015, 28(5): 182-185.
- [13] 肖永华, 王世东, 李靖, 等. 吕仁和辨治糖尿病医案症状、证候和治则解析 [J]. 北京中医药大学学报, 2011, 34(8): 560-564.
- [14] 王世东, 肖永华, 刘晓峰, 等. 吕仁和教授辨治糖尿病医案用药规律分析 [J]. 现代中医临床, 2016, 23(1): 33-37, 50.
- [15] 孙明月, 高蕊. 魏子孝治疗甲状腺疾病医案挖掘分析 [J]. 中国中医药信息杂志, 2012, 19(2): 25-27.
- [16] 吴中平, 何新慧, 柯雪帆. 医案查询统计分析系统介绍 [J]. 上海中医药杂志, 2003, 37(3): 54-56.
- [17] 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.
- [18] 任廷革, 刘晓峰, 高剑波, 等. “中医药基础数据库系统”介绍 [J]. 中国中医药信息杂志, 2001, 8(11): 90-91.
- [19] 李国辉. 脾胃湿热证与气象要素相关性的数据挖掘初步研究 [J]. 长沙: 湖南中医药大学, 2011: 7.
- [20] 李仕进, 陈蓉, 田玲, 等. 基于贝叶斯方法的中医“症-证”分析 [J]. 计算机工程, 2008(1): 212-214.
- [21] 黄晓红. 历代医家治疗足三阴虚损组方思想初探 [J]. 北京: 北京中医药大学, 2006: 3.
- [22] 徐玲艳. 名老中医沈炎南学术思想及医案整理 [J]. 广州: 广州中医药大学, 2013: 39.
- [23] 刘蕊. 许润三名老中医治疗排卵障碍性疾病的数据挖掘结果分析 [J]. 北京: 北京中医药大学, 2011: 5.
- [24] 许润三. 四逆散加味治疗输卵管阻塞 115 例总结报告 [J]. 中医杂志, 1987, 28(9): 41-42.
- [25] 郑志鹏. 干祖望诊治慢性咽炎的数据挖掘分析及用药规律的研究 [J]. 广州: 广州中医药大学, 2016: 3.
- [26] 王东方, 干千, 干祖望. 慢性咽炎咽干证候探析 [J]. 辽宁中医杂志, 1997, 24(7): 19-20.
- [27] 张慧. 清代情志相关性病证医案研究 [J]. 北京: 北京中医药大学, 2008: 38.
- [28] 凌雅睿. 基于数据挖掘方法的张书信教授治疗慢性便秘经验研究 [J]. 北京: 北京中医药大学, 2016: 38-39.
- [29] 申颖. 基于医案的名老中医学术思想及临证经验传承方法探索 [J]. 北京: 北京中医药大学, 2016: 20.
- [30] 宋洪娟, 李宛真, 王米渠. 论个案研究的思路拓展 [J]. 四川中医, 2011, 29(10): 52-53.
- [31] 刘晓峰, 任廷革, 高全泉, 等. 中医处方智能分析系统的研究与实践 [J]. 中国中医药信息杂志, 2007, 14(10): 97-99.
- [32] 黄苗, 肖永华, 刘晓峰, 等. 基于中医处方智能分析系统的吕仁和治疗慢性肾炎 1 例量化研究 [J]. 中国中医药信息杂志, 2016, 23(10): 42-45.

Review of Analysis Methods of Chinese Medical Records

HE Yancheng, XIAO Yonghua, YAN Pu, CHEN Jingshu, WANG Shidong

(Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing 100700)

ABSTRACT Chinese medical records are the most representative records of doctors' academic thinking and syndrome differentiation. Through the review of the literature on medical records analysis in recent years, based on the knowledge characteristics of explicit knowledge and tacit knowledge in traditional Chinese medical records, and the data characteristics of medical records and data that are difficult to standardize, starting from the status of research from group cases and individual case studies, combined with the advantages of individualized syndrome differentiation and treatment of traditional Chinese medicine, the necessity and feasibility of case data analysis was proposed.

Keywords Chinese medicine case; group case analysis; individual case analysis; tacit knowledge; traditional Chinese medicine intelligent prescription analysis system

(收稿日期: 2017-12-10; 修回日期: 2018-02-10)

编辑: 黄 健