

DOI: 10.13288/j.11-2166/r.2016.14.019

综 述

数据挖掘技术在针灸研究中的应用与展望

张小珊¹, 方剑乔^{2*}

(1. 贵阳中医学院针灸推拿系, 贵州省贵阳市市东路 50 号, 550002; 2. 浙江中医药大学)

[摘要] 通过回顾近 10 年数据挖掘技术在针灸领域的应用状况, 了解到针灸文献的数据挖掘研究多集中于临床常见疾病用穴规律分析、单个穴位应用研究、针灸研究数据库的构建、经穴特异性的证明、针刺手法的量化五个方面。基于以上研究成果, 结合数据挖掘在针灸研究中存在的数据整理、鉴别等方面的问题, 认为除了针灸文献数据的整理挖掘以外, 还要尽可能收集针灸临床诊治详细信息以获取大数据, 在运用数据挖掘技术探寻内在规律的同时还可确定个性化的最佳针灸治疗方案、提供更人性化的针灸治疗服务, 这对促进针灸技术的规范与发展具有重要意义。

[关键词] 数据挖掘; 针灸研究; 经穴特异性; 针刺手法; 针灸数据库

数据挖掘 (data mining) 是从大量数据中提取事先不知道的、动态的实际数据, 挖掘出未知的、有价值的模式或规律等知识的复杂过程, 它是一个多学科交叉的研究与应用领域, 涉及数据库技术、机器学习、模式识别、人工智能以及管理信息系统等各学科的成果^[1]。目前数据挖掘技术在不同领域取得了进展, 针灸领域中的许多学者也开始采用数据挖掘技术来探寻针灸治病的内在规律, 并已取得了一些重要的研究成果。本文回顾近 10 年数据挖掘技术在针灸领域的研究文献, 旨在为针灸研究未来发展探索新的途径。

1 数据挖掘技术在针灸研究中的应用与问题

1.1 应用现状

针灸学源远流长, 具有独特的理论体系和思维模式, 几千年来指导着临床实践, 积累了丰富的治疗经验。中医文献浩如烟海, 针灸文献和其他中医文献一样复杂多面、信息量难以估计。目前基于数据挖掘探索针灸治疗疾病内在规律的研究多集中于几个方面, 包括临床常见疾病用穴规律分析、单个穴位应用研究、针灸研究数据库的构建、经穴特异性的证明、针刺手法的量化等。其中对临床常见疾病用穴规律的研究最多, 如梁爽等^[2]收集中国知网数据库中近 20 年运用针灸治疗急性痛风性关节炎的临床文献, 建立针灸处方的主穴数据库, 采用

SPSS 17.0 和 Clementine 12.0 软件对数据进行描述性分析、关联规则分析和聚类分析, 从大量数据中分析出针灸治疗急性痛风性关节炎的临床选穴和配穴规律; 吴粮葶等^[3]通过采集文献建立数据库, 规范病症、腧穴名称, 统计中风后遗症所属病症的针灸病症谱及系统分类概况, 总结针灸治疗中风后遗症的腧穴谱, 结果发现, 中风后遗症的现代针灸文献病症谱分布呈偏向性, 腧穴谱遍布十四经脉, 首选阳经腧穴; 孙晶等^[4]通过对古代与三叉神经痛有关的针灸治疗文献的检索, 建立处方数据库, 运用针灸数据挖掘系统软件进行分析, 结果显示, 其选穴集中在头面部, 特定穴中交会穴使用最多, 腧穴配伍中“风池-率谷”使用频次最高。其他文献采用数据挖掘技术研究的病种还有偏头痛、糖尿病、冠心病等^[5-10]达 28 种。也有学者^[11]结合循证医学实践方法和中医文献学方法, 从古代医籍、现代期刊和专著、行业标准等信息源中采集古今针灸研究证据, 构建针灸研究数据库。此外, 有学者^[12-13]针对国外研究者对腧穴特异性的质疑, 运用数据挖掘技术整理、分析有关经穴效应特异性的古代文献和现代研究成果, 证明了经穴特异性的存在。数据挖掘在针刺量化中的应用方面, 杨霖等^[14]在针刺手法测定仪上进行 200 多次针刺手法操作, 取得大量针刺手法参数, 建立数据库, 采用 WEKA 数据挖掘软件, 运用关联规则和分类规则对其进行数据挖掘, 为针刺手法的量化做了有益探索。

* 通讯作者: fangjianqiao7532@163.com

1.2 存在的问题

数据挖掘的基础是“数据”，从目前的文献报道来看，用于研究的数据来源主要有古代文献记载、现代著作及期刊、开展实验获得 3 种形式。

就古代文献来说，针灸文献达三百余种，由于语言习惯的变迁、文字的不断演化、传抄过程中的错误、古籍的严重损毁等原因，给文献的阅读和整理带来巨大困难，对文献研究人员提出了更高要求，而我们目前的针灸文献整理虽然有了一些进展，但结果还不尽如人意。

就现代著作及期刊而言，随着社会的进步、信息的不断丰富，与针灸相关的现代文献不断增多，包括全国名家医案收集整理的著作和数以万计的科研论文。在这些论著中，有针灸名家数十载的经验总结，有科研工作者的研究成果，有临床医生的观察和发现，构成了针灸学宝贵的数据资源。但由于多方面原因，在这些文献中也存在着一些抄袭之作以及捏造的数据结果，阅读者需要花费更多的时间去辨别真伪，大大降低了研究效率，也使人们对数据挖掘结果存疑，从而使研究结果无法对临床产生指导意义。如果数据本身存在错误，即使采用最先进的数据挖掘技术，得到的结果也是错误的，但不具备任何参考价值，甚至可能会误导临床决策。

还有部分数据来源于研究者自己开展实验获取的数据，这样的资源大多是对针刺手法及其量效关系的研究，以多学科交叉的方式，与针灸以外的其他学科人员协作完成，为针刺手法的标准化研究做出了一定贡献。但目前这类研究尚处于初级阶段，研究者还在探索可供推广的研究方法，寻找适宜的数据挖掘技术，其研究成果的适用性及可重复性还有待临床验证。这需要广大针灸工作者与数据挖掘、数据管理人员及其他学科研究人员通力合作，共同致力于针灸这门传统医学的研究，而这样一个庞大团队的建立尚待各方努力。

2 数据挖掘技术在针灸研究中的应用展望

数据挖掘给针灸研究发展带来前所未有的契机，但作为一种技术，它仅仅是我们借以研究的一种工具，针灸领域的科学研究关键还是在于对针灸疗法本身的思考。不管是临床观察还是基础研究，最终目的都是为了更好地指导临床治疗。针灸临床以脏腑经络学说为指导，以辨证论治为法则，运用多种刺灸方法治疗各科疾病，在几千年的实践与探索中已形成了一套行之有效的诊疗体系。近年来，有人质疑经穴的存在，认为经穴没有特异性，面对

质疑之声，针灸研究者做出了强有力地回应，但同时我们也在思考，针灸领域确实有太多无法解答的疑问，比如经络腧穴的实质、针灸治病的机理、针刺手法的量化、各种复杂因素与疗效之间的关系等。面对诸多悬而未决的问题，我们认为，数据挖掘技术在针灸领域还能有助于解决以下问题。

2.1 确定个性化的最佳针灸治疗方案

完整的针灸治疗方案包含腧穴、刺灸方法、频次及疗程。首先，确定腧穴。现有的数据挖掘结果多是对腧穴或配穴出现频次进行筛选，虽然可以说明此穴在治疗该病症时被广泛应用，却不能说明此穴在治疗中是否起到了主要的治疗作用，且临床实际工作中往往多穴配伍，所以此类研究还不足以指导临床。针灸取穴固然可依据一定的法则，但是同一病症不同的医生或者同一个医生都可能出现多种选穴，究竟哪一组取穴疗效更好，疗程更短，更不易复发，更适合哪一类人群呢？根据现有的文献，我们还无从得知针对某一类患者取哪一个（组）腧穴疗效会更优；或者假设疗效无差异，如何确定哪一个（组）腧穴更能缩短病程、更能减少复发，甚至是让患者感觉更加舒适。其次，确定刺灸方法。以治疗腰痛为例，假设选用了阿是穴，可选择的方法有针刺、艾灸、温针灸、拔罐、刺血、穴位注射等，临床医生会根据患者的年龄、性别、体质、病情等选择一种或几种方法联合应用，但对某一个体使用哪种方法最优，临床医生即使有多年的经验也仍然未必有十足的把握。最后，确定频次和疗程。针灸治疗的频次即每次针灸干预的间隔时间。针灸临床观察类文献中经常出现这样的描述，针刺“每日 1 次”“隔日 1 次”“5 次为 1 个疗程”“10 次为 1 个疗程，间隔 2 日再进行下 1 个疗程”等，虽然间隔时间不统一，但临床疗效都很肯定。面对这样的现象便会产生如下问题：每日 1 次与隔日 1 次针灸干预的疗效是否一样；每日 1 次与隔日 1 次各适宜于什么情况；如何确定 1 个疗程的干预次数等。对于这些问题，从古代文献到现代研究都很难直接找到明确的答案。也有学者对此做了专门的研究，查阅大量文献，对古今文献中针刺频次的论述做出总结，并结合现代时间生物学对针刺频次影响针灸疗效的机理做出论述，得出结论是：针刺治疗的频次不仅要根据疾病的不同性质、不同时期灵活掌握，还应考虑到患者的年龄、体质等各种因素；并希望形成标准的最佳治疗频次指南对临床工作进行指导，更好地提高疗效，为患者节约时间和开支^[15]。由此可见，针灸频次的确定并不容易，

其原因在于: 1) 针灸干预方法多样, 涉及到针刺、灸法、拔罐、放血、刮痧等, 每一种方法的频次选择不尽相同。2) 患者个体差异包括年龄、体质、病情、心理、社会角色不同, 能够承受的针灸干预强度也不一样。如临床中个别住院患者因无法承受每天针刺而要求出院到门诊行隔日 1 次的针灸治疗, 以减轻针刺的痛苦及恐惧感。3) 患者在医院的诊疗情况各异。就目前情况来看, 部分住院患者的治疗频次可由医生依据患者病情来确定, 而部分门诊患者则大多依据医生的门诊时间来确定。4) 患者根据自己的经济状况和行程安排来选择接受治疗的频次。总体来说, 针灸频次与刺灸方法、患者、医生及其他社会因素有关, 其中的内在联系尚未有定论。

如果能够真实、详尽地收集到临床资料, 包括患者的一般情况、四诊资料、辨证论治过程、配穴处方、刺灸方法、对每个症状的改善情况、疗程、疗效、随访, 以及医生的资历等数据, 建立数据库, 用数据挖掘技术处理这些海量的、不清晰的、看似无关联的数据, 找出其中的规律, 我们将有望在尊重针灸临床医疗现状的同时, 分析出腧穴与病症之间、刺灸方法与病症之间、频次及疗程与病症之间、患病人群与病症之间的关联性, 发现其中的内在联系及规律, 从而得出针对患者个体的某一病症的最佳针灸治疗方案。

2.2 提供更人性化的针灸治疗服务

既往受到研究方法的限制, 针灸医生关注的重点是疗效, 这一点本身是毋庸置疑的, 但作为针灸疗法中最常用的针刺, 其治疗注重“得气”与否, 而“得气”与否关系到针灸治疗的疗效和预后, 依赖于医生和患者的主观感受。临床工作中, 针灸医生非常重视患者的感觉, 但在病例记录及文献报道中却鲜有提及, 这就导致临床工作中极为重要的信息没有得到很好的整理和利用。我们认为, 除了患者疾病本身的改变外, 患者的主观感受是应该详实记录的主要内容。患者在接受针灸治疗过程中会告知症状的转变情况、针刺后是否得气、治疗方法的舒适度、对施术者的好恶等信息, 如果能对这些信息加以数据挖掘, 有可能找出疾病发生、发展与患者、治疗方案、疗效、耐受度甚至是舒适程度等多因素间的相互关联, 其分析结果将有利于实现“以患者为中心”, 指导医生制定一套疗效最好的、患者更易于接受的、更能够坚持治疗的方案。

综上, 通过对针灸临床详细信息的收集、整理并采用数据挖掘技术进行挖掘, 寻找临床实际工作

中复杂因素间的内部联系, 将会提供给医生最佳的针灸治疗指南, 亦会给予患者更多的人文关怀, 从而使针灸方法更好地为患者服务。我们有理由相信, 在鼓励学科交叉的大背景下, 从针灸自身发展的需要出发, 借助多学科的技术和方法, 针灸学科的发展方向将逐渐明晰, 发展道路将更加宽广。

参考文献

- [1] 王小妮. 数据挖掘技术 [M]. 北京: 北京航空航天大学出版社, 2014: 2.
- [2] 梁爽, 黄凯裕, 许岳亭, 等. 基于数据挖掘的针灸治疗急性痛风性关节炎临床选穴规律分析 [J]. 世界科学技术-中医药现代化, 2014, 16(12): 2598-2604.
- [3] 吴粮葶, 李瑛, 任玉兰. 基于数据挖掘针灸治疗中风后遗症的病症谱及腧穴谱 [J]. 中国针灸, 2013, 33(6): 547-552.
- [4] 孙晶, 方剑乔, 陈勤, 等. 与三叉神经痛相关病症的古代针灸处方规律分析 [J]. 上海针灸杂志, 2013, 32(1): 66-68.
- [5] 陈裕. 当代名中医针灸治疗偏头痛用穴特点的数据挖掘 [J]. 深圳中西医结合杂志, 2012, 22(3): 148-150.
- [6] 张航飞, 邓宏勇, 赵玲, 等. 针灸治疗糖尿病数据挖掘分析 [J]. 上海针灸杂志, 2014, 33(12): 1173-1177.
- [7] 陈霞, 卢圣锋, 朱冰梅, 等. 针灸治疗冠心病临床选穴规律的数据挖掘研究 [J]. 南京中医药大学学报, 2014, 30(5): 417-421.
- [8] 贾红玲, 张学伟, 张永臣. 数据挖掘技术在针灸治疗腰痛古代文献中的应用 [J]. 针灸临床杂志, 2013, 29(11): 40-45.
- [9] 李享, 寿依夏, 任玉兰, 等. 古代不同时期针灸治疗眩晕用穴特点的数据挖掘研究 [J]. 中国针灸, 2014, 34(5): 511-515.
- [10] 杨洁, 任玉兰, 吴曦, 等. 基于数据挖掘技术的针灸治疗贝尔面瘫 RCT 文献的用穴规律分析 [J]. 中华中医药杂志, 2010, 25(3): 348-351.
- [11] 任玉兰, 梁繁荣, 吴曦, 等. 基于数据挖掘的针灸临床循证决策支持系统研究 [J]. 中华中医药杂志, 2011, 26(4): 795-797.
- [12] 梁繁荣, 曾芳, 赵凌, 等. 经穴效应特异性及其基本规律 [J]. 中国针灸, 2009, 29(2): 129-132.
- [13] 任玉兰, 赵凌, 陈勤, 等. 数据挖掘技术在经穴选用及其特异性研究中的应用 [J]. 中医杂志, 2010, 51(1): 47-51.
- [14] 杨霖, 洪菲, 杨华元. 针刺手法数据挖掘的关联规则与分类 [J]. 上海针灸杂志, 2014, 33(11): 1064-1070.
- [15] 钟玉涛, 黄东来, 高邈. 针刺频次对针灸疗效影响的探索 [J]. 辽宁中医药大学学报, 2012, 14(10): 132-133.

(收稿日期: 2015-09-16; 修回日期: 2015-10-27)

[编辑: 崔京艳]