

基于 CiteSpace 的门诊大数据研究趋势及热点可视化分析

徐静娟,朱 枫,刘华联

(常州市第一人民医院门诊办公室,江苏 常州 213003)

摘要:目的 基于 CiteSpace 软件对门诊大数据研究的相关文献进行可视化分析,分析该领域的研究热点和趋势,为今后研究工作开展提供借鉴。方法 在中国知网数据库(CNKI)中,以“大数据”OR“big data”AND“门诊”为主题词,检索从 2013 年 1 月—2023 年 6 月的文献。运用 CiteSpace 对文献的发表时间、作者、机构、期刊、关键词进行可视化分析。结果 ①共纳入 224 篇文献,每年发文量呈稳步上升趋势,2018 和 2020 年达到最高;②研究主要在信息和管理类的期刊发表;③作者合作网络分析显示,黄绍中、陈纯、黄宇飞、谢小平、张勇发文量较多,陈纯、张勇、谢小平、黄绍中已经形成了稳定的研究团队;④研究机构主要集中在大学及大学附属医疗机构,各团体间的合作并不明显。发文量较多的前 5 所机构分别为上海交通大学、电子科技大学、福建省医疗保险管理中心、华中科技大学、北京大学第三医院;⑤关键词聚类分析后,形成 10 个聚类群,分别为大数据、数据挖掘、智慧医疗、疫情防控、门诊、医院、医疗费用、机器学习、管理、电子病历;⑥关键词突现分析显示,研究趋势为数据挖掘、机器学习、传染病防控、门诊电子病历和智慧门诊。结论 大数据已成为门诊管理的研究热点之一,作者、机构之间的协作还需要加强,未来仍需扩展门诊大数据应用场景。

关键词:大数据;门诊管理;CiteSpace;可视化分析

中图分类号:R197.3

文献标识码:A

DOI:10.3969/j.issn.1006-1959.2024.14.008

文章编号:1006-1959(2024)14-0043-06

Research Trend and Hot Spot Visualization Analysis of Outpatient Big Data Based on CiteSpace

XU Jing-juan,ZHU Feng,LIU Hua-lian

(Outpatient Department of Changzhou First People's Hospital,Changzhou 213003,Jiangsu,China)

Abstract: **Objective** To visually analyze the related literature of outpatient big data research based on CiteSpace, analyze the research hotspots and trends in this field, and provide reference for future research work. **Methods** The literature from January 2013 to June 2023 were retrieved from CNKI database with the keywords of 'big data' AND 'outpatient'. CiteSpace was used to visualize and analyze the publication time, authors, institutions, journals, and keywords of the literature. **Results** ①A total of 224 literature were included, and the number of articles published per year showed a steady upward trend, reaching the highest in 2018 and 2020. ②The research literature were mainly published in the journals of information and management. ③The analysis of authors' collaborative network showed that HUANG Shao-zhong, CHEN Chun, HUANG Yu-fei, XIE Xiao-ping, and ZHANG Yong had more publications, and that CHEN Chun, ZHANG Yong, XIE Xiao-ping, and HUANG Shao-zhong had formed a stable research team. ④The research institutions were mainly concentrated in the universities and university-affiliated medical institutions, and cooperation among groups was not obvious. The top five institutions with the largest number of publications were Shanghai Jiao Tong University, University of Electronic Science and Technology of China, Fujian Medical Insurance Management Center, Huazhong University of Science and Technology, and Peking University Third Hospital. ⑤According to keyword clustering analysis, 10 clusters were formed, which were big data, data mining, intelligent medical healthcare, epidemic prevention and control, outpatient, hospital, medical cost, machine learning, management, and electronic medical records. ⑥According to keyword emergence analysis, the research trends were data mining, machine learning, infectious disease prevention and control, outpatient electronic medical records and intelligent outpatient service. **Conclusion** Big data has become one of the research hotspots in outpatient service management, and the collaboration between authors and institutes needs to be strengthened. In the future, it is still necessary to expand the application scenarios of outpatient big data.

Key words: Big data; Outpatient management; CiteSpace; Visualization analysis

门诊是大多数患者及家属进入医院的第一站,门诊人员密集度高、流动性大,且存在多重工作人

员、职能、流程、环节和场所的交错,门诊服务质量直接反映了医院的管理水平和技术实力,进而影响到患者及家属就诊的体验感和满意度^[1]。门诊信息系统,包括医院信息系统(hospital information system, HIS)、医学影像存档与通讯系统(picture archiving and communication systems, PACS)、实验室信息管理系统(laboratory information management system, LIS)、电子病历(electronic medical record, EMR)和

作者简介:徐静娟(1976.11-),女,江苏常州人,硕士,主任护师,主要从事医院行政管理工作

通讯作者:刘华联(1983.10-),男,安徽颍上县人,硕士,主任医师,主要从事卫生行政管理工作

收费系统,储存着大量的数据信息^[2]。对这些门诊大数据进行智能挖掘和分析,发掘大数据中蕴含的知识与规律,可为管理者的决策提供有效支撑^[3]。CiteSpace 软件是应用于文献计量学研究的重要工具之一,是在科学计量学、数据和信息可视化背景下发展起来的一款引文分析软件,通过可视化的手段来呈现文献的知识结构、规律和分布情况^[4]。本研究基于 CiteSpace 软件,分析近 10 年内关于大数据在门诊管理应用研究的相关文献,分析门诊大数据研究现状、研究热点和研究趋势,为门诊大数据的深入研究提供依据。

1 资料与方法

1.1 数据来源 数据来源于中国知网(CNKI)。检索方式为高级检索中的“主题”检索;检索策略:“大数据”OR“big data”AND“门诊”;时间:2013 年 1 月 1 日–2023 年 6 月 30 日。纳入标准:与门诊大数据相关的期刊论文、学位论文。排除标准:新闻报道、征稿启示、会议论文、科普宣传、重复发表及与主题不相关的文献。2 名课题组成员分别进行文献检索,阅读文题、摘要及全文进行文献筛选,意见不一致时与第 3 名研究者共同讨论决定。

1.2 数据转换 将 CNKI 中所纳入文献,选择 Refworks 形式导入相应文件夹,并以 download_XXX 形式命名,文献的信息包括发表年份、作者、机构、期刊、题名、摘要、关键词等。

1.3 软件参数设置 使用的软件版本为 CiteSpace 6.2.R2,生成相应的可视化图谱。CiteSpace 分析的参数设置:时间跨度(time slicing)选择 2013–2023 年,

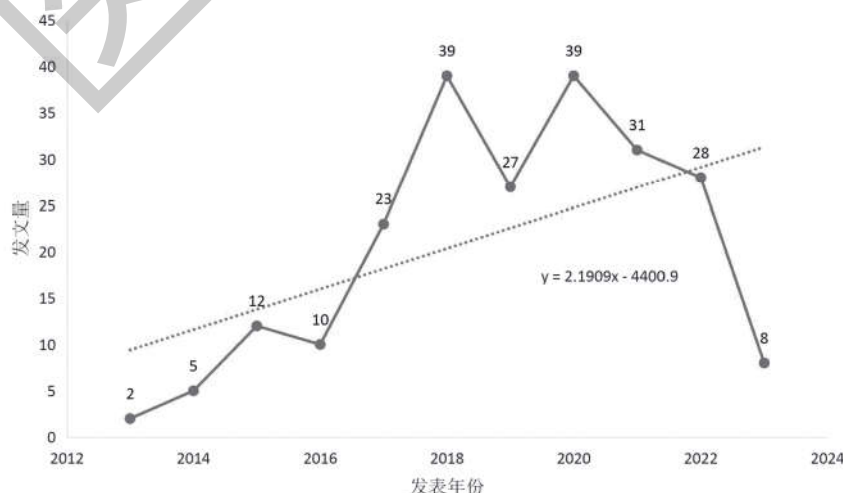
时间切片为每 1 年;术语来源 (Term Source)为 Title、Abstract、Author、Keywords 和 Keywords Plus;节点类型 (Node Types) 分别为 Author、Institution、Keyword;联结强度 (Links Strength)为 Cosine;阈值为 Top 50;裁剪方式 (Pruning)为 Pathfinder、Pruning sliced works 和 Pruning the merged network;可视化 (Visualization) 为 Cluster View –Static 和 Show Merged Network。用 Excel 2018 进行频数计算。

2 结果

2.1 发文量的时间分布 共搜索出文献 280 篇,根据纳入及排除标准,阅读题目及摘要删除 45 篇(研究主题不符 27 篇、研究类型不符 18 篇),阅读全文排除文献 11 篇(研究主题不符),共排除 56 篇。从年度发文量看,2013–2023 年发文量呈逐渐上升趋势,发文量回归方程为 $Y=2.1909X-4400.9$ 。2013–2016 年间发文量较少,每年不超过 15 篇。2017–2018 年发文量大幅增长,2018–2022 年发文量保持平稳,每年在 20 篇左右,2023 年仅纳入了 6 个月的文献,发文量较少,见图 1。

2.2 期刊发文量分析 发文量前 10 名期刊共发表门诊大数据应用相关文献 59 篇,其中《中国数字医学》发文量最多(11 篇),见表 1。

2.3 作者合作网络分析 作者合作网络分析得到 706 个节点,1496 根连线,网络密度为 0.006。合作网络显示,黄绍中、陈纯、黄宇飞、谢小平、张勇发文量较多,陈纯、张勇、谢小平、黄绍中已经形成了稳定的研究团队,其他大部分作者分布分散,发文量较少,见图 2。



注:虚线为文献线性趋势线;2023 年发文量截至 6 月 30 日。

图 1 发文量分布图

表 1 2013–2023 年发文量前 10 名期刊情况[n(%)]

期刊名称	发文量	期刊名称	发文量
《中国数字医学》	11(18.64)	《医院管理论坛》	5(8.47)
《中国医疗保险》	8(13.56)	《电子科技大学》	4(6.78)
《中国卫生信息管理杂志》	8(13.56)	《江苏卫生事业管理》	4(6.78)
《现代医院管理》	6(10.17)	《上海交通大学》	4(6.78)
《医学信息学杂志》	5(8.47)	《中国医院》	4(6.78)



图 2 发文作者合作网络分析图

2.4 机构合作网络分析 机构合作网络分析得到 347 个节点,354 根连线,网络密度为 0.0059。合作网络显示,众多机构对门诊大数据的应用进行了深入研究,主要集中在大学及大学附属医疗机构,各团体间的合作并不明显。发文量较多的前 5 所机构机构分别为上海交通大学(9 篇)、电子科技大学(4 篇)、福建省医疗保险管理中心(3 篇)、华中科技大学(3 篇)、北京大学第三医院(3 篇),见图 3。

2.5 关键词分析

2.5.1 关键词共现分析 关键词进行共现分析得到 417 个节点,680 根连线,网络密度为 0.0078。关键词研究热点主要集中在大数据、门诊、数据挖掘、医院、医疗费用、疫情防控、机器学习、分级诊疗、智慧

医疗、医院管理、信息化,见图 4、表 2。

2.5.2 关键词聚类分析 对关键词进行聚类分析得到关键词聚类图,共形成一致性较高的 10 个聚类,分别为大数据、数据挖掘、智慧医疗、疫情防控、门诊、医院、医疗费用、机器学习、管理、电子病历,见图 5。

2.5.3 关键词突现分析 关键词突现表现为在短时期内出现频率的明显增加,可代表该领域的研究趋势^[9]。本研究的突现关键词共 15 个,按突现强度由高到低排列,依次为医院、分级诊疗、疫情防控、数据挖掘、机器学习、传染病、人工智能、医疗数据、区域医疗、数据分析、管理、预约诊疗、智慧门诊、电子病历和信息化,见图 6。



图 3 机构合作网络分析图



图 4 关键词共现分析图

表 2 关键词中心性和频次(前 10 位)

位次	关键词	频次	中心性	位次	关键词	频次	中心性
1	大数据	52	0.59	6	医疗费用	8	0.07
2	Big data	15	0.02	7	疫情防控	6	0.13
3	门诊	14	0.07	8	机器学习	6	0.03
4	数据挖掘	14	0.14	9	分级诊疗	6	0.03
5	医院	14	0.17	10	智慧医疗	6	0.04

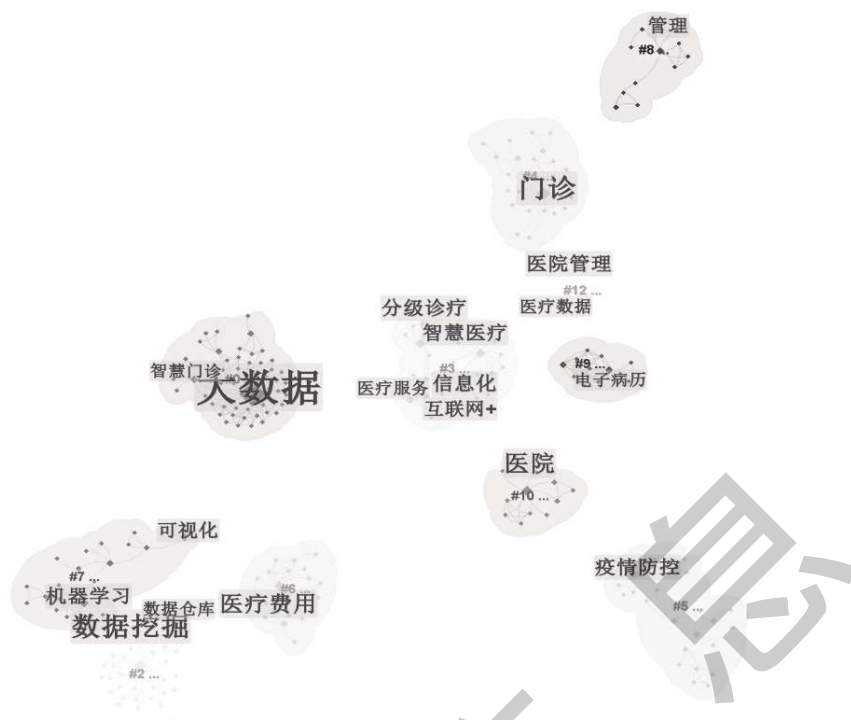


图 5 关键词聚类分析图

Top 15 Keywords with the Strongest Citation Bursts

Keywords	Year	Strength	Begin	End	2013 - 2023
分级诊疗	2017	1.81	2017	2018	<div></div>
医疗数据	2018	0.95	2018	2019	<div></div>
区域医疗	2018	0.82	2018	2019	<div></div>
预约诊疗	2018	0.46	2018	2020	<div></div>
人工智能	2019	1.21	2019	2020	<div></div>
管理	2016	0.63	2019	2020	<div></div>
信息化	2017	0.34	2019	2020	<div></div>
医院	2015	2.13	2020	2021	<div></div>
疫情防控	2020	1.65	2020	2023	<div></div>
数据分析	2020	0.78	2020	2021	<div></div>
数据挖掘	2015	1.54	2021	2023	<div></div>
机器学习	2019	1.35	2021	2023	<div></div>
传染病	2021	1.28	2021	2023	<div></div>
电子病历	2014	0.43	2021	2023	<div></div>
智慧门诊	2018	0.43	2021	2023	<div></div>

图 6 关键词突现图

3 讨论

3.1 门诊大数据已成为门诊管理的研究方向之一
本研究结果显示,2013–2023 年门诊大数据相关研究的文献数量逐年提升,2018 和 2020 年呈现峰值。信息和管理类的期刊对门诊大数据研究有持续的关注。作者和研究机构共现网络分析显示,陈纯、张勇、谢小平、黄绍中的研究团队较稳定,其他不同作者及研究机构之间联系不紧密,未形成明显的合作网络,

表明本领域内合作需进一步加强。

3.2 门诊大数据的应用

3.2.1 门诊信息共享 收集门诊的数据信息,建立疾病、专科检查、临床术语、诊断和知识数据库及质量管理体系可以实现大数据框架下的同质化疾病管理^[6,7]。通过门诊大数据,可建立门诊诊疗信息共享平台,实现 HIS、LIS、PACS、EMR 等信息数据的共享^[8],也有助于医院和基层医疗机构之间的分级

诊疗工作的实施^[9]。

3.2.2 门诊流程优化 基于门诊就诊各个环节的大数据分析,如号源数量、门诊人流量、等候时间、医疗人力数量等,实现门诊医疗资源的合理分配,改进门诊服务流程,可提高门诊服务水平,改善门诊患者就医体验^[2,10,11]。

3.2.3 疫情防控管理 传染病的疫情防控是医院的感染管理和公共卫生的重要工作内容,建立基于大数据的传染病信息化监测预警和多点触发机制,有助于提高传染病的流调效率,传染病的早发现、早诊断、控制疾病传播^[12,13]。

3.2.4 疾病风险预警 疾病风险预测模型与门诊健康大数据管理平台相结合,利用机器学习、数据挖掘、深度学习等数据处理方式^[14-16],建立疾病预警机制,对门诊患者的个人健康状况进行评估与风险预测^[17-19]。

3.2.5 门诊费用监控 通过数据仓库导出门诊费用相关数据,可以分析门诊费用情况及其影响因素,针对性采用控费^[20,21]。采用异常数据挖掘技术,对门诊收费异常点进行排查,建立预警机制,可为门诊医生用药、检查和治疗提供参考,也可减少财务异常费用结算发生^[22]。

3.3 门诊大数据的研究趋势 根据文献关键词聚类 and 突现的综合分析,今后门诊大数据的研究趋势集中在数据挖掘、机器学习、传染病防控、门诊电子病历和智慧门诊,门诊管理者可进一步在这些方面拓展门诊大数据的应用场景。

综上所述,门诊大数据的研究热度呈逐年上升的趋势,不同研究者和研究机构之间的协作性和关联度还需要进一步加强。门诊大数据在门诊信息共享、门诊流程优化、疫情防控、疾病风险预警、门诊费用监控等门诊管理环节中发挥了积极的作用。未来,门诊管理者仍需进一步拓展门诊大数据的应用场景,从而提高门诊管理效率和服务质量,改善门诊患者就医体验。

参考文献:

- [1]王平.《医疗机构门诊质量管理暂行规定》意义解读[J].中国卫生质量管理,2023,30(2):20-23.
- [2]刘刚,李娟.基于大数据分析的门诊就诊影响因素及诊疗流程动态优化研究[J].成都医学院学报,2022,17(1):110-113.
- [3]吴琨,郑欣丽,朱嘉辉,等.基于 PCA-GRA-BK 算法的医疗大数据分析[J].华中师范大学学报(自然科学版),2023,57(3):

364-372.

- [4]陈超美,陈悦,侯剑华,等.CiteSpace II:科学文献中新趋势与新动态的识别与可视化[J].情报学报,2009,28(3):401-421.
- [5]陶紫晶,陈格格,崔玉容,等.基于 VOSviewer 和 CiteSpace 可视化分析“脑-肠轴”理论的研究热点与前沿[J].世界中西医结合杂志,2023,18(5):876-886.
- [6]刘瑾.眼科结构化门诊电子病历大数据管理与临床应用[D].汕头:汕头大学,2021.
- [7]陈柯羽,唐武军,于洁.信息化在中医医院门诊病历质量控制中的应用及 SWOT 分析[J].中国医药导报,2021,18(4):155-158.
- [8]张彦彬.基于超融合架构的电子病历信息共享平台研究[D].北京:北京交通大学,2021.
- [9]周利宏.以门诊电子病历为入口的分级诊疗建设[J].解放军医院管理杂志,2018,25(7):685-687.
- [10]王其军,李扬,单清,等.大数据下的专科专家门诊供需精细化管理研究[J].江苏卫生事业管理,2021,32(7):907-910.
- [11]黄昊,赵平,罗贤斌,等.信息化提升门诊服务能力[J].中国数字医学,2016,11(10):26-28.
- [12]陈媛,丘美娇,陈宝,等.综合医院发热门诊信息化建设对传染病监测预警及登记报告的影响[J].海南医学,2021,32(17):2293-2297.
- [13]周莉莉,余洋.基于大数据平台的院内传染病流调系统设计与实现[J].中国数字医学,2023,18(4):68-72.
- [14]王能才,王玉珍,张海英,等.基于人工智能的医疗大数据中心设计与构建[J].中国医学装备,2022,19(2):1-5.
- [15]崔晓旭.基于数据挖掘的疾病预测组合模型研究[D].北京:北京交通大学,2019.
- [16]安莹,黄能军,杨荣,等.基于深度学习的心血管疾病风险预测模型[J].中国医学物理学杂志,2019,36(9):1103-1112.
- [17]张习梅,杨露,南原.疾病预警在健康大数据管理平台中的应用[J].医学信息学杂志,2021,42(2):49-52,65.
- [18]相静,王玖,胡西厚.健康医疗大数据驱动下的疾病风险评估与预测方法探析[J].中国卫生信息管理杂志,2018,15(3):329-333.
- [19]郭佳凯.基于数据仓库的电子病历大数据分析[D].沈阳:中国医科大学,2022.
- [20]陈纯,张勇,黄绍中,等.2005-2014 年福建省直单位参保人群中心脑血管疾病患者门诊就诊机构及费用分布的概况及分析[J].中国全科医学,2017,20(24):3008-3014.
- [21]程小恩,温川飙,许强,等.基于大数据分析的中医门诊医疗诊费点阵模型构建[J].贵州中医药大学学报,2020,42(4):93-98.
- [22]李会敏,王亚林,李涛.基于医疗信息化的医院门诊转账结算风险控制方法研究[J].自动化与仪器仪表,2020(5):146-149.

收稿日期:2023-08-24;修回日期:2023-09-04

编辑/杜帆