

·数字中医药·

本文引用:刘琦,韦昌法,刘东波,杨嘉诚,周婉珠,匡文轩,晏峻峰.袁肇凯教授学术研究可视化知识图谱分析[J].湖南中医药大学学报,2019,39(4):489-494.

袁肇凯教授学术研究可视化知识图谱分析

刘琦,韦昌法,刘东波,杨嘉诚,周婉珠,匡文轩,晏峻峰*
(湖南中医药大学信息科学与工程学院,湖南长沙 410208)

[摘要] **目的** 借助文献计量工具对袁肇凯教授学术研究进行可视化图谱分析。**方法** 在中国知网(CNKI)上以“袁肇凯”为关键词进行作者检索,检索得文献 289 篇,借助 CiteSpace 进行文献的可视化分析,分别从作者发文量、合作作者网络、合作机构网络、突现词检测、关键词共现网络进行分析。**结果** 袁教授的主要研究热点为:冠心病、养心通脉片、证候及其本质研究、心血瘀阻证、动脉粥样硬化作用、骨髓间充质干细胞、早发冠心病。主要研究方向为冠心病血瘀证及其治疗。与袁教授合作密切作者依次为黄献平、胡志希、孙贵香、郑景辉、王东生、谢梦洲、李杰。**结论** 袁教授紧紧围绕冠心病研究领域不断向外拓展,积极运用光电容积仪、人工神经网络等新方法,从系统生物学、代谢组学、基因组学等多角度进行研究,其治学专而精,富有开拓进取的创新精神。

[关键词] 袁肇凯;文献可视化分析;冠心病;血瘀证;CiteSpace

[中图分类号]R249

[文献标志码]A

[文章编号]doi:10.3969/j.issn.1674-070X.2019.04.012

Professor YUAN Zhaokai's Knowledge Map of Academic Thought and Visual Analysis

LIU Qi, WEI Changfa, LIU Dongbo, YANG Jiacheng, ZHOU Wanzhu, KUANG Wenxuan, YAN Junfeng

(School of Information Science and Engineering, Hunan University of Chinese Medicine, Changsha, Hunan 410208, China)

[Abstract] **Objective** To visualize Professor YUAN Zhaokai's academic thoughts by means of bibliometric tools. **Methods** A total of 289 papers were retrieved by using "YUAN Zhaokai" as the key word in the author's search on Chinese National Knowledge Infrastructure (CNKI). Visual analysis of the literature was carried out by CiteSpace. The author's publication number, author's cooperative network, institutional cooperative network, salient word detection and keyword co-occurrence network were analyzed respectively. **Results** Professor YUAN's main research hotspots were from Qi and blood syndrome differentiation of hepatic fibrosis of schistosomiasis in the early ages, to effects of Yangxin Tongmai Tablets on insulin resistance syndrome and atherosclerosis, to the study on the essence of syndrome, to the study on blood stasis syndrome of coronary heart disease and bone marrow mesenchymal stem cells in recent years. The authors who cooperated closely with Professor YUAN were HUANG Xianping, HU Zhixi, SUN Guixiang, ZHENG Jinghui, WANG Dongsheng, XIE Mengzhou and LI Jie. **Conclusion** Professor YUAN has been expanding the research field of coronary heart disease, actively using new methods such as photovoltaic volumetric instrument and artificial neural network, studied it from the perspectives of system biology, metabonomics, genomics and so on. He is expert in science and has a pioneering spirit of innovation.

[Keywords] YUAN Zhaokai; visual analysis of literature; coronary heart disease; heart blood stasis syndrome; CiteSpace

[收稿日期]2018-12-16

[基金项目]湖南省研究生创新课题(CX2018B479)。

[作者简介]刘琦,男,在读博士研究生,主要从事中医药信息学研究。

[通讯作者]* 晏峻峰,女,教授,博士研究生导师,E-mail:teacheryan@qq.com。

袁肇凯教授为首届国家级中医药名师,湖南中医药大学教授,师承全国首批名老中医专家郭振球教授,始终耕耘在中医药临床治疗、科研、教学一线。发表了诸多的学术论文,但面对庞大的文献数量,以阅读、归纳法进行分析存在一定的局限性、主观性和片面性,因此,对袁教授的学术论文相关数据进行可视化分析,能够明晰地展现其学术脉络的流动及变迁^[1]。本文借助 CiteSpaceV.5.0.R4 对中国知网上有记录以来到 2018 年 12 月 15 日止袁肇凯教授的学术论文进行分析、挖掘、图形呈现来揭示其研究主题演变的过程,通过阶段划分和关键文献的解读分析,探索学者个人知识图谱的演变规律。

1 方法和数据

本文使用的研究工具是采用美国德雷克塞尔大学信息科学与技术学院陈超美教授基于 JAVA 语言开发的文献数据挖掘和可视化软件 CiteSpace^[1],版本号为 2.R1.3.9.2018。它主要基于共引分析理论和寻径网络算法等,对特定领域文献进行计量分析,以寻找出科学领域演化的关键联结路径和知识转折点,并融合了聚类分析、社会网络分析等多种分析方法,绘制可视化图谱来形成对学科演化潜在动力机制的分析和学科发展趋势的探测。本文运用该软件对 CNKI(中国知网)上以“袁肇凯”为作者的相关文献进行包括作者发文量,合作机构、关键词、突现词等分析。

1.1 文献数据检索策略

以 CNKI 数据库为文献数据来源,采取作者检索的方式进行检索。一般来说,查准率高,查全率就低;查全率高,查准率就低^[2]。陈超美教授认为应该优先保证查全率^[3],数据的完备性比数据的准确性更为重要。不相干的检索数据在后续进行科学知识图谱分析中,会自动暴露出来。

因此,为保证查全率,以一窥袁教授学术研究全貌,设定检索关键词为“袁肇凯”,检索方式为作者检索,跨库选取了:期刊、教育期刊、特色期刊、博士、硕士、国内会议、国际会议、报纸、年鉴、学术辑刊。文献分类目录:全选。不限定检索方式为第一作者或通讯作者等,亦不限定作者的单位机构。以袁教授为第一作者或通讯作者的文献是直接与袁教授学术

领域相关,而以袁教授为普通作者的文献可以认为与袁教授团队相关,有充足的理由认为其与袁教授的研究领域相关。在检索的过程中发现,与袁肇凯教授同名的学者在 CNKI 数据库中并不存在,也就是在所检索的 339 篇文献中并未发现与袁教授不相关的领域或团队的文献出现,有充足的理由认为 339 篇文献中均与袁教授相关。检索时间跨度为有记录以来到 2018 年 10 月 10 日。

按以上检索策略将检索所得的 339 篇文献选中,经筛选掉重复文献后得到了 289 篇文献,按 ref-works 格式导出,并以 CiteSpace 所要求的格式命名。

1.2 文献数据预处理

运用 Notepad 7.5 代码编辑器清洗文献数据,合并同义词,标准化文献数据格式。保存后利用 CiteSpace 的“data”文献数据格式转换功能转换文献记录格式,使其成为 CiteSpace 可以分析的数据。

1.3 节点设置

由于检索到的文献出版时间跨度为 1982 至 2017 年,故在 CiteSpace 中将时间切片(Time Slicing)设置为 from 1982 to 2017,每一切片间隔(#years Per Slice)设置为 3 年,文献数据量较小,网络裁剪为 None,其他的为默认值。

2 文献可视化结果

CNKI 数据库并未开放文章引文数据^[4],故本次分析主要以合作网络和共现网络的分析为主。本研究主要围绕“袁肇凯”关键作者作可视化分析,国家地区合作不作分析。

2.1 发文量分析

图 1 为作者论文发表量的分析。横坐标为发文的年代,纵坐标为发文量。图中可见作者前期发文较少,到 1992 年开始发文量不断增加,2020 年迅速攀升,2002 年到 2013 年间形成发文量小高潮。

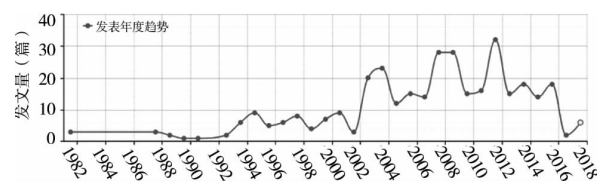


图 1 不同年度袁肇凯教授发文量分析

2.2 合作作者网络分析

结点类型(Node Types)设置为:Author,阈值设

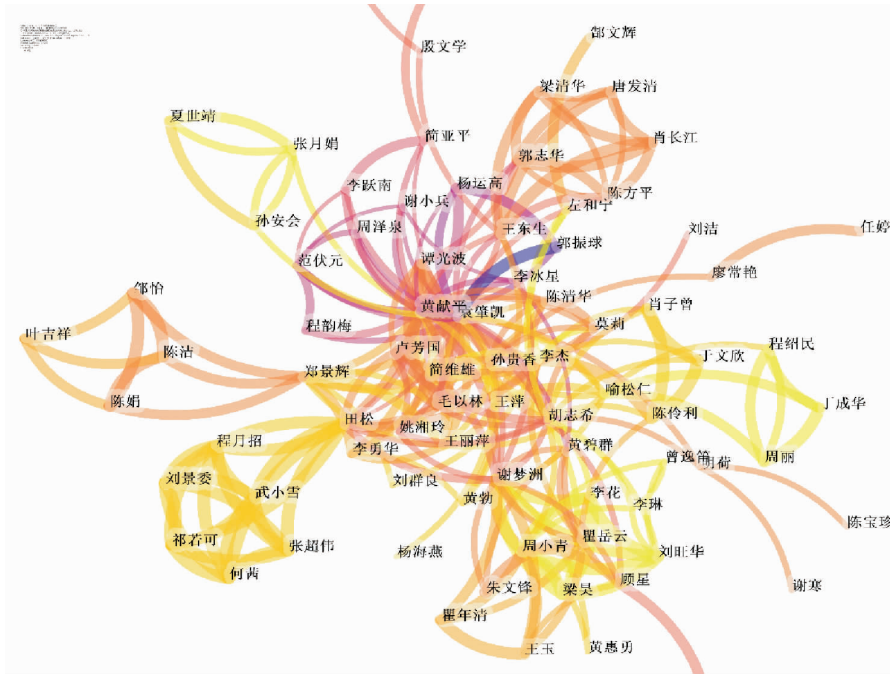


图 2 与袁肇凯教授合作作者网络分析

置为:Top N%为 50%。得到初步的可视化知识图谱(图 2)。其中连线颜色代表的是首次合作的时间,颜色的冷、暖色调代表的是距今的远、近。而网络中节点位置为与袁肇凯教授合作的密切程度。中心性前十位分别是:袁肇凯(0.58)、黄献平(0.39)、胡志希(0.39)、谢梦洲(0.28)、田松(0.26)、王东生(0.23)、郑景辉(0.20)、孙贵香(0.17)、喻松仁(0.14)、李杰(0.13)。合作频率排名前十位的是:黄献平、胡志希、简维雄、孙贵香、郑景辉、王东生、王丽萍、谢梦洲、李杰、陈清华。

2.3 合作机构网络分析

结点类型(Node Types)设置为关键词 Institution, 阈值设置为:Top N%为 100%。得到机构合作网络图谱,隐去袁教授所在机构“湖南中医药大学”

节点,选取时间线视图(图 3)进行可视化分析。其中连线代表机构合作关系, 连线颜色代表的是首次合作的时间,颜色的冷、暖色调代表的是距今的远、近。节点大小为机构在发文记录中出现次数,节点位置为活跃的时间,处于靠左的为比较久远的年代,靠右的为较近的年代。

图 3 显示:与袁教授合作较多的机构依次为:江西中医药大学,中南大学湘雅,广西中医药大学附属瑞康医院。

2.4 关键词共现网络分析

结点类型(Node Types)设置为关键词 Keyword, 阈值设置为:Top N%为 30%。得到关键词共现图谱如图 4。其中连线颜色代表的是首次共现的时间,颜色的冷、暖色调代表的是距今的远、近。网络中节点



图 3 与袁肇凯教授合作机构网络分析

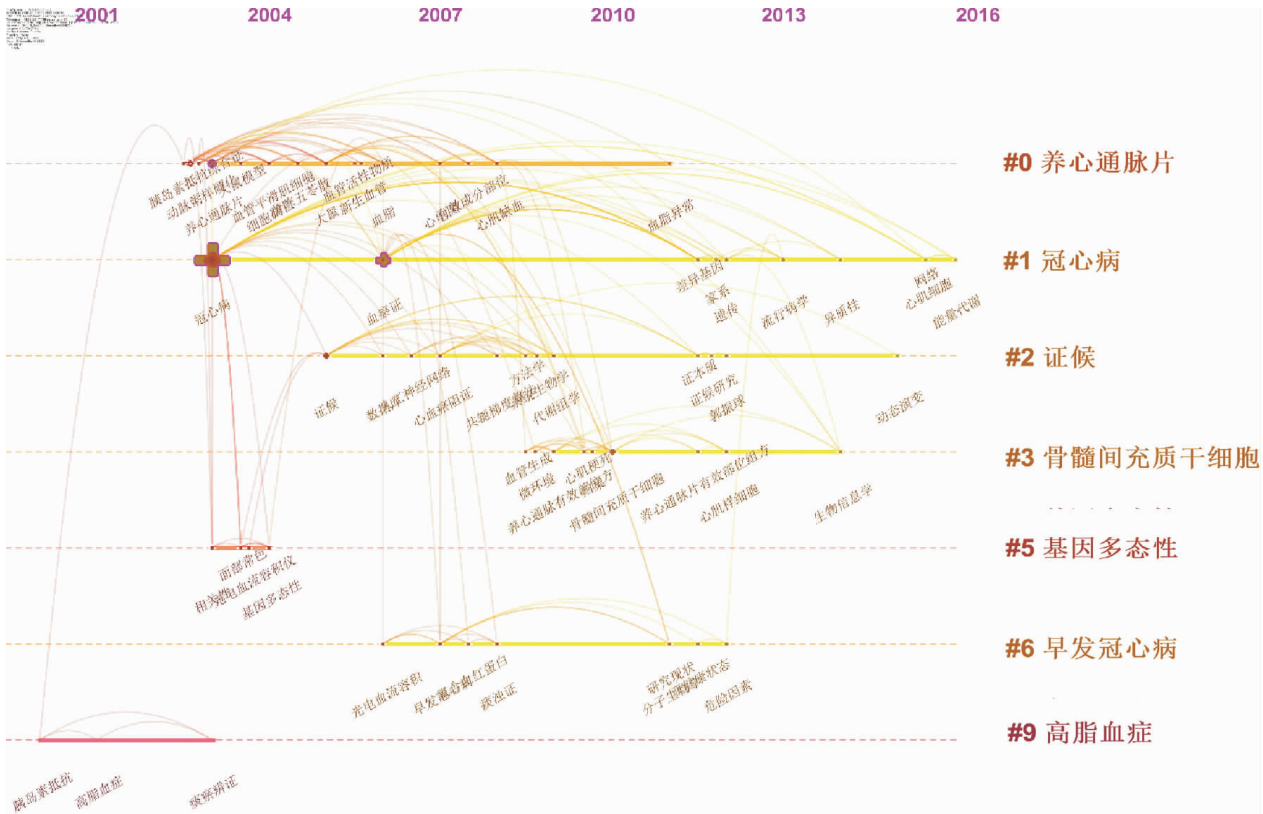


图 6 袁教授学术研究的关键词聚类时间线视图

表 1 关键词聚类信息

聚类 ID	节点数	剪切值(S)	标签(LLR)
#0	20	0.739	养心通脉片、胰岛素抵抗综合征、动脉粥样硬化、大鼠
#1	14	0.831	冠心病 血瘀证
#2	14	0.759	证候、代谢组学、心血瘀阻证
#3	12	0.82	骨髓间充质干细胞,养心通脉有效部位方
#5	8	0.8	基因多态性
#6	8	0.896	早发冠心病,分子生物学
#9	3	0.975	高脂血症,痰瘀辨证

性来看,袁教授主要关注在:冠心病、养心通脉片、血瘀证、心血瘀阻证、证候、骨髓间充质干细胞、大鼠。可以看到这些词同时具有高频性。

从聚类视图的分析来看,袁教授团队涉及到的领域可大致分为 7 个研究方向,对表 1 的分析看出聚类#0 主要聚焦在养心通脉片的疗效上,或对胰岛素抵抗综合征、动脉粥样硬化作用,可能涉及到了动物实验,对该方向的研究从 2003 年持续到了 2011 年;聚类 #1 主要关注的是冠心病的相关研究,大量地跟血瘀证相关,该聚类持续时间最长,从 2004 年一直持续到了 2017 年,同时剪切值和包含的关键词数相对较高,说明此方向为袁教授团队 2004 年之后的主

要研究方向;聚类 #2 为研究证候本质,试图用代谢组学进行解释,从心血瘀阻证方面作为突破口,该研究方向从 2006 年持续到了 2015 年;聚类 #3 关注的方向是骨髓间充质干细胞,并研究养心通脉有效部位方对其的影响,时间跨度为从 2010 年到 2014 年;聚类 #5 以探讨基因多态性为主,该聚类较小,年代较久远且持续时间最短;聚类 #6 聚焦在亚健康状态“早发冠心病”相关的生物指标上面,该聚类活跃在 2008 年到 2014 年之间;聚类 #9 关注在高脂血症及其痰瘀辨证方向,该聚类最小、年代也最久远,同时剪切值最高,是为与其他方向关联度最小的一个聚类,可以认为这是早年袁教授的一个研究方向,与之后的研究相对相关度不大。

3.2 研究趋势

以上的结果可以较明显地看出袁教授学术研究的演变路线。从发文量上看,可以看出袁教授在知网上有记录的学术论文发表时间跨度为 1982 至 2017 年,学术影响力持续了 36 年。1994 年之前年发文量约 2 篇;1994 年到 2002 年年发文量浮动在 6 篇左右;到了 2003 年开始文献发表量呈爆发式增长,达到了 23 篇;2003-2016 年间,有三次年发文量处于 24

篇以上的波峰,分别是:2003-2004年、2008-2009年、2012年,其他年间均维持在一个高位的年14篇文献量的平稳期。

对关键词突现性分析需要结合发文量分析来研究袁教授的研究趋势。从图5可知1994年之前袁教授的研究尚未足够形成聚焦点,到1994-1996年间出现了“血吸虫病肝纤维化”,对应的正是1994-1996年之间发文量的一个小高峰;1998年和2001年2012年的两个小高峰对应的正是“气血辨证”;2003-2004年、2003-2006年、2003-2007年形成的波峰与“胰岛素抵抗综合征”“养心通脉片”“动脉粥样硬化”研究大量出现有关;2008-2009年波峰的形成与“证候”相关研究的大量出现有关;2010-2014年之间出现的最高发文量波峰则与“骨髓间充质干细胞”的研究相关。2013年之后到2017年袁教授团队则主要转向了对“冠心病血瘀证”相关的研究。

3.3 社会合作关系

从袁教授合作网络的连线颜色分析来看,袁教授早年的合作网络较小,未处于中心位置,可知早年袁教授的合作作者之后未见有明显合作,可能与关注焦点变更有关。合作网络整体颜色丰富,代表各个时段袁教授均有较高频的合作作者,是为袁教授科研教学两不误,培养了众多的“合作作者”,呈现桃李满天下之势。而具有高中心性和高合作关系的是:黄献平、胡志希、孙贵香、郑景辉、王东生、谢梦洲、李杰。表示这些人在以袁教授为中心的合作关系网中具有较大影响力和较好的团队领衔作用。

在机构合作网络中可得出,袁教授早期(2004年)与中南大学有过短暂合作,后期(2012年)与江西中医药大学有过大量的合作。其中时间跨度较长的是湖南中医药大学中医诊断学国家重点学科,从2010~2013年出现较为明显,需要结合袁教授实际工作经历进一步分析。

4 小结

综合来看,袁教授团队研究趋势从早年的血吸虫病肝纤维化的气血辨证,到养心通脉片及其对胰岛素抵抗综合征、动脉粥样硬化的作用,到对证候本质的研究,再到近年的冠心病血瘀证、骨髓间充质干细胞研究。主要研究方向为冠心病血瘀证及其治疗。

纵观袁教授36年的学术发文跨度,袁教授紧紧围绕心血管系统领域展开研究^[7],同时以主要方向冠心病研究领域为基础不断向外拓展,从光电容积仪、系统生物学、代谢组学、基因多态性、骨髓间充质干细胞、分子生物学、人工神经网络等新工具、新方法和新的视角进行研究,启示后来科研工作者进行研究需要保持专而精的精神,紧紧围绕自己熟悉的领域开展研究,同时保持开拓进取的创新精神,不断探索研究的外围领域,积极善用新的研究方法或工具从事自己的研究。

参考文献

- [1] 陈悦,陈超美,刘则渊,等.CiteSpace知识图谱的方法论功能[J].科学学研究,2015,33(2):242-253.
- [2] 王阳.基于超像素的图像分割算法研究[D].兰州:兰州理工大学,2018.
- [3] 蒲杨.叙词在INIS数据库检索中的应用研究[J].图书情报工作,2015,59(S2):171-176.
- [4] 相德宝,张弛.议题、变迁与网络:中国国际传播研究三十年知识图谱分析[J].现代传播(中国传媒大学学报),2018,40(8):73-77.
- [5] 潘黎,王素.近十年来教育研究的热点领域和前沿主题——基于八种教育学期刊2000-2009年刊载文献关键词共现知识图谱的计量分析[J].教育研究,2011,32(2):47-53.
- [6] 乐思诗,王平尧,应贤军.科学计量学在技术预见研究中的应用——以新材料研究为例[J].图书馆研究,2016,46(2):118-122.
- [7] 刘吉勇,袁肇凯.袁肇凯教授辨证治疗心病经验[J].湖南中医药大学学报,2017,37(3):281-284.

(本文编辑 李路丹)