基于 ESI 的临床医学学科热门论文分布及 其研究热点

白 蓓 蒲 剑 曹 玥 刘梦菡 董珊珊

(四川大学华西医院 成都 610041)

[摘要] 选用 ESI 数据库,对临床医学热门论文国家及期刊分布进行文献计量分析,对引用排名前 62 的论文进行共被引分析。结果表明,美国和《新英格兰医学杂志》在临床医学领域的热门论文最多。心血管疾病的预防和控制、肿瘤、丙型肝炎以及心血管疾病药物临床试验是研究热点。

[关键词] ESI: 临床医学: 热门论文: 计量分析: 研究热点

[中图分类号] R-056 [文献标识码] A [**DOI**] 10. 3969/j. issn. 1673 - 6036. 2016. 05. 015

ESI – based Analysis of the Distribution and Research Highlights of Hot Papers on Clinical Medicine BAI Bei, PU Jian, CAO Yue, LIU Meng – han, DONG Shan – shan, West China Hospital, Sichuan University, Chengdu 610041, China

[Abstract] Based on ESI database, the paper makes bibliometric analysis of the distribution of hot papers on clinical medicine in countries and periodicals. As indicated by the results, the United State and the *New England Journal of Medicine* contain the largest number of hot papers in the field of clinical medicine. The research highlights are the prevention and control of cardiovascular diseases and tumors, hepatitis C and cardiovascular diseases.

[Keywords] ESI; Clinical medicine; Hot papers; Bibliometric analysis; Research focus

1 引言

基本科学指标数据库(Essential Science Indicators, ESI)是衡量国家、研究机构、研究人员研究绩效,掌握学科发展趋势和动向的重要文献计量及评价工具^[1]。ESI 共分为 22 个学科,临床医学(Clinical Medicine)是其中之一。ESI 从研究领域、作者、研究机构、期刊、国家、研究前沿等角度提供顶尖论文(top paper)、高被引论文(highly cited

papers)、热门论文(hot papers)的索引信息。热门论文是指某学科领域发表在最近两年间,同时在最近两个月内被引次数排在前 0.1%以内的论文。本研究通过分析临床医学学科热门论文,对引用频次、通讯作者所在国家、发表期刊进行统计,分析排名靠前期刊及国家的热门论文引用频次有无差异,对排名前 62(被引频次 > 150)的论文进行共被引分析,聚类关系紧密的论文,揭示当前临床医学学科的研究前沿,为临床医学领域学者提供学科发展趋势、研究热点等相关信息。

[收稿日期] 2016-01-18

[作者简介] 白蓓,实习研究员;通讯作者:蒲剑,副研究员。

2 资料及方法

2.1 热门论文国家及发表期刊数据

数据检索日期为2015年11月4日,收集以下信息:引用频次、通讯作者所在国家、发表期刊。 ESI中收录的论文1篇可能有多个参与国家,该篇 论文被计入所有参与国家热门论文数中,因此本 研究按照通讯作者所在国家对论文国家来源重新 进行梳理,识别出不同国家热门论文的引用频次 分布。

2.2 共被引分析

选取引用频次超过 150 次的热门论文,排除 4 篇重复论文,最终有 58 篇论文纳入共被引分析。通过 Web of Science (WOS) 的论文检索功能,在检索历史中将重复论文的引用频次进行合并,得到该篇论文引用频次合集,将该合集纳入共被引分析。同理,将 58 篇论文的检索结果两两合并,得到论文两两之间的共被引次数。

2.3 统计分析

本研究采用文献计量法、描述性统计、假设检验对热门论文期刊与国家分布进行分析。论文的共被引分析通过聚类分析和多维尺度分析实现。使用 SPSS 统计软件包(版本 22)完成相关数据分析。

3 结果

3.1 热门论文频次及国家、期刊分布

ESI 共检索出热门论文 506 篇,排除重复论文 12 篇(仅保留引用频次高的重复论文),共计 494 篇论文最终纳入统计。494 篇热门论文中,引用频次 1 000 次以上的有 1 篇 (0.20%),500~999 次的有 4 篇 (0.81%),300~499 的有 6 篇

(1.21%), 200~299的有19篇(3.85%), 100~199的有86篇(17.41%), 100次以下的有378篇(76.5%)。

由于热门论文引用频次数据属于偏态分布资 料,所以用中位数描述集中趋势,四分位数间距描 述离散趋势,国家热门论文引用情况,见表1。期 刊热门论文引用情况,见表2。本研究对排名前10 国家及期刊引用频次分布分别进行了两两秩和检 验,见表3、表4。美国、法国、意大利、德国、澳 大利亚与荷兰热门论文的引用频次分布有统计学差 异 (P<0.05), 其他国家两两间没有统计学差异。 结果显示,荷兰的热门论文引用频次中位数较其他 国家小,论文排名相对靠后,而美国、法国、意大 利、德国、澳大利亚的热门论文引用频次中位数相 对较高,有更大的影响力。《新英格兰医学杂志》 与《柳叶刀肿瘤学》、《临床肿瘤杂志》的引用频次 分布有统计学差异 (P < 0.05), 其他期刊两两间没 有统计学差异,说明《新英格兰医学杂志》比这两 个期刊有更大的影响力。

表 1 热门论文国家引用频次分布

国家	热门论文量	引用频次中位数	百分比		
	(篇)	(四分位数间距)	(%)		
美国	253	54(73.5)	51. 21		
英国	42	44(62)	8. 50		
法国	25	55(93)	5.06		
意大利	24	77.5(95)	4. 86		
德国	23	57(68)	4. 66		
澳大利亚	20	56(61.25)	4. 05		
加拿大	20	42. 5 (73. 5)	4. 05		
荷兰	15	21 (43)	3.04		
丹麦	8	31 (48. 25)	1.62		
中国	8	60.5(121.5)	1.62		
其他国家*	56	71(112)	11. 34		
全部国家*	494	53.5(74)	100.00		

^{*}包括无国别学术团体。

表 2 热	门论文期刊引	用频次分布
-------	--------	-------

排名	期刊名称	热门论文量 (篇)	引用频次中位数 (四分位数间距)	百分比(%)
1	《新英格兰医学杂志》	147	68 (93)	29. 76
2	《柳叶刀》	52	60 (84.5)	10. 53
3	《美国医学会杂志》	22	55. 5 (59. 5)	4. 45
4	《柳叶刀肿瘤学》	20	45 (55)	4. 05
5	《临床肿瘤杂志》	17	35(65)	3. 44
6	《美国心脏学院杂志》	16	46.5(65)	3. 24
7	《欧洲心脏病杂志》	16	36(86.5)	3. 24
8	《自然》	12	65. 5 (74. 75)	2. 43
9	《循环》	11	61 (114)	2. 23
10	《血液》	10	59 (47. 25)	2. 02
11	其他期刊	171	38(46)	34. 62
_	全部期刊	494	53.5(74)	100.00

表 3 排名前 10 国家频次分布两两秩和检验结果 (P 值)

	美国	英国	法国	意大利	德国	澳大利亚	加拿大	荷兰	丹麦	中国
美国	-	0.300	0. 915	0. 164	0.517	0.809	0.466	0.008 *	0. 302	0. 816
英国	0.300	-	0.460	0.065	0. 199	0. 572	0.851	0.079	0.507	0.846
法国	0. 915	0.460	-	0.401	0.788	0. 723	0.493	0.028 *	0.352	0.918
意大利	0. 164	0.065	0.401	-	0. 537	0. 225	0. 128	0.003 *	0.070	0. 535
德国	0.517	0. 199	0. 788	0. 537	-	0.488	0. 242	0.007 *	0. 172	0.674
澳大利亚	0.809	0. 572	0.723	0. 225	0.488	-	0.698	0.036 *	0.354	1.000
加拿大	0.466	0.851	0.493	0. 128	0. 242	0.698	-	0.064	0.601	0. 98
荷兰	0.008 *	0.079	0.028 *	0.003 *	0.007 *	0.036 *	0.064	-	0.466	0. 238
丹麦	0.302	0.507	0. 352	0.070	0. 172	0. 354	0.601	0.466	-	0.645
中国	0.816	0.846	0.918	0. 535	0.674	1.000	0.98	0. 238	0.645	_

^{*}P < 0.05,有统计学差异。

表 4 排名前 10 期刊频次分布两两秩和检验结果 (P 值)

	《新英格兰	《柳叶刀》	《美国医	《柳叶刀	《临床	《美国心脏	《欧洲心脏病	// 44 凸 //	《循环》	// m >床 \\	
	医学杂志》		学会杂志》	肿瘤学》	肿瘤杂志》	学院杂志》	杂志》	《自然》	《 1/目 J L //	《血液》	
《新英格兰		0. 153	0. 360	0.009 *	0.046 *	0. 202	0.074	0. 922	0.889	0. 299	
医学杂志》	_										
《柳叶刀》	0. 153	-	0. 939	0. 123	0. 296	0.734	0. 347	0. 475	0. 593	0.752	
《美国医学	0.360	0. 939		0. 217	0. 279	0.672	0. 298	0.488	0.807	0.734	
会杂志》			_								
《柳叶刀	0.009 *	0. 123	0. 217		0.964	0.479	0. 937	0.070	0. 227	0. 373	
肿瘤学》				_							
《临床肿瘤	0.046 *	0. 296	0. 279	0. 964		0. 444	0. 929	0. 180	0. 225	0.570	
杂志》					_						
《美国心脏	0. 202	0.734	0. 672	0.479	0. 444		0.642	0.302	0. 544	1.000	
学院杂志》						_					
《欧洲心脏病	ā 0.074	0. 347	0. 298	0. 937	0. 929	0.642	-	0. 159	0. 294	0. 421	
杂志》											
《自然》	0. 922	0. 475	0.488	0.070	0. 180	0.302	0. 159	-	0.786	0. 539	
《循环》	0.889	0. 593	0.807	0. 227	0. 225	0. 544	0. 294	0. 786	-	0.605	
《血液》	0. 299	0.752	0.734	0. 373	0.570	1.000	0. 421	0. 539	0.605	_	

^{*}P < 0.05,有统计学差异。

3.2 热门论文共被引分析

共被引的定义是两篇不同论文同时被其他论文 引用的次数,共被引次数越多,说明这两篇论文的 关系越紧密^[2]。通过对论文进行共被引分析,可以 探索某学科领域特定国家、研究人员、研究机构的 活跃程度,还可以为特定学科领域提供前沿发展趋 势[3]。论文共被引分析主要分为3个步骤,构建原 始共被引矩阵,转化原始矩阵为相关矩阵,多元统 计分析[4-5]。58 篇论文两两间共被引结果被构建为 原始共被引矩阵,图1列出了部分原始共被引矩 阵。图 1 中 A1, A2 ·····, A58 代表论文题目(论 文题目对应编号详见附录)。本研究删去了与其他 论文之间关系不紧密的论文, 以矩阵中非零项少于 5 的论文为标准[6],最后剩余 42 篇热门论文,论文 编号未改变。利用 SPSS 将 42×42 共被引原始矩阵 转换为相关系数矩阵,见图2。将相关系数矩阵作 为数据源进行聚类分析和多维尺度分析。聚类结果 如图 3 所示, 多维尺度分析, 见图 4。

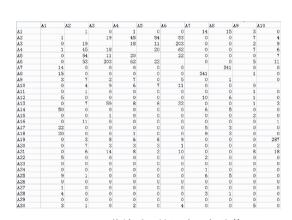


图 1 58×58 共被引原始矩阵 (部分截图)

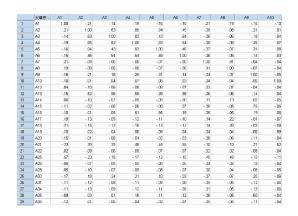


图 2 42×42 相关系数矩阵 (部分截图)

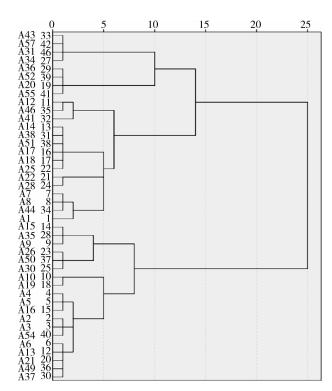


图 3 聚类分析结果

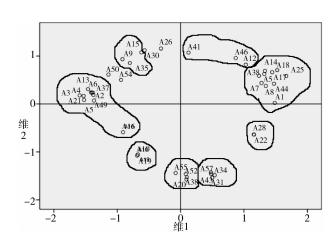


图 4 多维尺度分析结果

聚类分析将相同研究领域论文进行聚类,多维尺度分析显示出相关研究领域之间关系的紧密程度^[5]。根据聚类分析和多维尺度分析结果,本文将42篇入选论文的研究领域分为以下6个方面。(1)心血管疾病(16篇)。治疗心血管疾病的研究4篇(A36、A52、A20、A55),包括药物治疗静脉血栓1篇,药物治疗房颤的临床试验、荟萃分析各1篇,华法林对人工瓣膜患者疗效的比较1篇。药物使用对2型糖尿病患者心血管结局的影响研究2篇

(A10、A19)。其他心血管疾病的研究 9 篇 (A4、 A5、A2、A3、A6、A13、A21、A49、A37),包括 高血压管理指南 2 篇,动脉粥样硬化、心力衰竭、 糖尿病/心血管疾病管理指南各1篇,心脏疾病及 卒中统计数据1篇,手术治疗高血压1篇,生活方 式干预对2型糖尿病患者心血管结局的影响研究1 篇,肥胖预防与控制1篇。温度管理对心脏骤停的 影响研究1篇(A16)。(2)除血液肿瘤外其他系 统肿瘤(12篇)。肿瘤治疗相关文章9篇(A14、 A38、A51、A17、A18、A25、A7、A8、A44),包 括药物治疗胰腺癌、肺癌、结直肠癌各1篇,放射 治疗骨转移前列腺癌 1 篇,癌症药物耐药性研究 1 篇,膀胱移行细胞癌分子特征研究1篇,药物治疗 黑色素瘤 2 篇, 药物治疗胶质母细胞瘤 1 篇。各类 癌症数据统计1篇(A1)。乳腺癌治疗指南2篇 (A22、A28)。(3) 丙型肝炎(4篇)。Sofosbuvir 合 并其他药物治疗丙型肝炎的临床试验3篇(A43、 A57、A31) 以及三联疗法治疗丙肝肝硬化的临床试 验 1 篇 (A34)。(4) 血液系统肿瘤(3 篇)。靶向 药物治疗慢性淋巴细胞白血病 1 篇 (A12), 靶向药 物治疗套细胞淋巴瘤 1 篇 (A46),慢性骨髓性白血 病综述 1 篇 (A41)。(5) 肥胖 (3 篇)。低收入及 中等收入国家孕产妇及儿童营养不良和超重研究1 篇 (A15), 儿童和成人肥胖患病率 2 篇 (A35、 A9)。(6)疾病负担(4篇)。美国和全球疾病负担 各1篇(A30、A50),精神和物质使用障碍引起的 疾病负担1篇(A26),慢性肾脏疾病综述1篇 (A54)。42 篇论文中, 临床试验有 21 篇 (药物临 床试验17篇),占50%,指南或综述类的论文有12 篇,占28.6%,提供统计数据的横断面研究有9 篇,占 21.4%。心血管疾病的预防和控制(7 篇),肿瘤药物临床试验(6篇),丙型肝炎药物临 床试验(4篇)以及心血管疾病药物临床试验(3 篇)是研究热点。42篇论文的研究内容、研究方向 以及研究类型。

4 讨论

4.1 国家及期刊分布

美国在热门论文数量上远领先于世界其他国

家,这与之前的研究结果一致^[6-8],说明美国近年来在临床医学领域始终保持领先水平。荷兰与美国、法国、意大利、澳大利亚热门论文频次分布两两秩和检验结果有统计学差异,可能是其肿瘤及心血管疾病研究领域的论文比例小,其他领域研究者不如肿瘤及心血管疾病领域的研究者多,故被引率相对较低。在临床医学领域,热门论文主要分布在影响因子较高的期刊中,这可能与影响因子的计算方式与期刊论文平均被引率有关,并且影响因子高的期刊更受高质量论文的青睐^[7]。《柳叶刀肿瘤学》、《临床肿瘤杂志》与《新英格兰医学杂志》引用频次分布有差异的原因可能是前两个期刊的论文收录范围更窄,研究方向更集中,故被引率相对较低,但仍受到专业内学者的重视。

4.2 研究热点

邱均平等人 2010 年的共被引分析显示肿瘤是 当时的研究热点[6],经过5年的时间,除肿瘤外, 心血管疾病以及丙型肝炎也成为了研究热点。肿瘤 和丙型肝炎的研究主要聚焦在疾病治疗的临床试验 上,而心血管疾病除了聚焦疾病治疗外,更加注重 疾病的管理与预防。心血管疾病具有发病率高、死 亡率高的特点,预防心血管疾病,控制危险因素如 高血压、糖尿病、肥胖尤为关键[9]。在共被引分析 的 42 篇论文中, 有 21 篇论文是药物或治疗方式的 临床试验, 其次是指南或综述, 再次是横断面研 究。同类疾病的药物临床试验共被引频次高,在多 维尺度分析中更集中, 如治疗丙型肝炎的药物研究 (A31、A34、A43、A57)、治疗黑色素瘤的药物研 究(A7、A8)、治疗心血管疾病的药物研究(A20、 A36、A52、A55)。指南性研究作为行业标准被引 频次高,与同领域研究的共被引频次也较高(A3、 A4、A5、A6)。横断面研究作为疾病数据支撑被引 频次高,但是部分论文的研究内容广泛,与其他论 文的共被引频次较低,例如排名第1的2014年美国 癌症数据统计以及疾病负担类的研究(A1、A30、 A50)。从聚类分析和多维尺度分析中可以看出,肥 胖、糖尿病和心血管疾病关系相对密切,可能与它 们是心血管疾病以及心血管不良事件发生的危险因

素有关[10-12]。

4.3 中国临床医学学科影响力

与2010年相比,中国临床医学学科的影响力有了很大提高,研究显示当时中国没有热门论文^[6],而如今已有8篇,更是跻身进入世界前10的行列。8篇论文中有研究禽流感病毒的论文2篇,肿瘤、心血管疾病、肥胖、呼吸系统疾病、消化系统疾病、感染性疾病各1篇。中国热门论文的分布频次与世界前9国家相比没有差异,说明中国热门论文的质量已与前9国家达到同一层次。近年来中国临床医学领域论文质量有较大提高可能与转化医学理念的深入,国家对临床医学研究的重视,以及国家对新发、突发疫情的快速响应有关^[13]。

4.4 研究局限性及未来方向

本研究只分析了临床医学领域大的研究方向, 缺乏对三级学科、亚专业的分析。在今后的工作 中,应根据不同专业科研人员需求,更加细化学科 研究方向,为科研人员提供该专业的学科发展动 态、发展前沿以及热点趋势。

5 结语

随着科学技术不断发展,特别是信息化技术的应用推广,数据作为重要资源提供各类信息、辅助各项决策^[14]。科研管理人员运用大数据资源优势,获得学科发展状况、发展趋势与研究热点,为科研人员提供主要研究领域相关信息,将为科研政策制定、科研发展方向选择提供科学、准确依据,为提高学科核心竞争力提供有力支持。

参考文献

1 熊璐. 美国《基本科学指标数据库》的研究与应用 [D]. 武汉: 武汉大学, 2005.

- 2 Small H. Co citation in the Scientific Literature: a new measure of the relationship between two documents [J]. Journal of the American Society for Information Science, 1973, 24 (4): 265, 269.
- 3 耿海英. 共引分析方法及其应用研究 [D]. 北京: 中国科学院研究生院(文献情报中心), 2007.
- 4 赵常志. 引文网状关系分析及其技术方法和步骤 [J]. 情报学报, 1992, 11 (5): 381-388.
- 5 邱均平,马瑞敏,李晔君.关于共被引分析方法的再认识和再思考 [J].情报学报,2008,27 (1):69-74.
- 6 邱均平,杨瑞仙.基于 ESI 的学科热门论文的计量研究——以临床医学学科为例 [J].情报科学,2010,1 (1):53-57.
- 7 陈汐敏,丁贵鹏,接雅俐,等.我国临床医学研究领域 论文产出状况国际比较研究[J]. 医学信息学杂志, 2013,34(7):45-50.
- 8 段文文, 缪幽竹. 基于 ESI 的中国临床医学学科影响力分析——以大陆地区高校和机构为例 [J]. 科技情报开发与经济, 2015, 25 (14): 142-143.
- 9 Ryden L, Grant P J, Anker S D, et al. ESC Guidelines on Diabetes, Pre – diabetes, and Cardiovascular Diseases Developed in Collaboration with the EASD: the Task Force on diabetes, pre – diabetes, and cardiovascular diseases of the European Society of Cardiology (ESC) and developed in collaboration with the European Association for the Study of Diabetes (EASD)[J]. Eur Heart J, 2013, 34 (39): 3035 – 3087.
- 10 何疆春,李田昌.心血管疾病风险评估的现状与展望[J].心血管病学进展,2013,34(1):50-55.
- 11 中国 2 型糖尿病防治指南 (2013 年版) [J]. 中国糖尿病杂志, 2014, 22 (8): 2-42.
- 12 陈丽,钱梦华,凌利民,等.上海市虹口区成人肥胖现 状与心脑血管疾病危险因素聚集性分析[J].中国初级 卫生保健,2015,29 (10):62-64.
- 13 许培扬. 中国疾病预防控制工作进展(2015年)[J]. 首都公共卫生, 2015, 9(3): 97-101.
- 14 丛培民. 从政策研究视角看大数据对科研管理与决策的 影响 [J]. 科研信息化技术与应用,2013,4(6):29-35.