

· 述评 · 综述 · 论坛 ·

药品不良反应信号挖掘文献计量学分析

丁月霞¹ 李硕硕² 郭晨煜¹ 尹爱群³ 唐启令¹

(1. 青岛大学附属烟台毓璜顶医院药学部 山东烟台 264000; 2. 烟台大学药学院;

3. 烟台市药品不良反应监测中心)

摘要 目的:深入了解药品不良反应信号的研究概况,客观反映相关国家、机构和科学家在研究药品不良反应的领域中具备的科学能力和影响力。**方法:**以科学引文索引(SCI)数据库 Web of Science 为检索平台,以“adverse reaction signal mining”“adverse reaction data mining”为主题词进行高级检索,采用文献计量学的方法,分析以不良反应信号挖掘为主题的相关文献在国家/地区、机构、作者来源、出版物、年发文量、年引文量和引用排名前 10 名的情况。同时,以中国知网(CNKI)中的中国期刊全文数据库为检索平台,高级检索“不良反应信号挖掘”“不良反应数据挖掘”为主题词的所有中文文献并进行文献计量学的分析。检索时间截止到 2018 年 6 月 30 日。**结果:**在 Web of Science 数据库中检索到英文文献 191 篇,年发文量与年引文数以 2013 为最多,h-index 为 36,美国和法国的发文量为 103 篇,占总数的 53.93%。在 CNKI 中检索到的中文文献为 41 篇,研究最多的机构为第二军医大学,发文量为 11 篇,占总数的 26.83%。**结论:**关于药品不良反应信号挖掘的研究发展迅速,且一直保持着一定的研究热度,但挖掘技术和数据利用的方法还需要进一步探索。

关键词 药品不良反应;信号挖掘;数据挖掘;文献计量学

中图分类号:R181.3⁺5 文献标识码:A 文章编号:1005-0698(2019)01-0054-05

Adverse Drug Reactions Signal Mining Biblio-metrological AnalysisDing Yuexia¹, Li Shuoshuo², Guo Chenyu¹, Yin Aiqun³, Tang Qiling¹

1. Department of Pharmacy, The Affiliated Yantai Yuhuangding Hospital of Qingdao University, Yantai 264000, Shandong, China; 2. School of Pharmacy, Yantai University; 3. Yantai Adverse Drug Reaction Monitoring Center

ABSTRACT Objective: To gain an in-depth understanding of the research profile of adverse reaction signals and to objectively reflect the scientific capabilities and influence of relevant countries, institutions, and scientists in the field of studying adverse drug reactions. **Methods:** Using the Science Citation Index (SCI) database Web of Science as a search platform, advanced search using “adverse reaction signal mining” “adverse reaction data mining” as the subject heading, and using bibliometric methods to analyze adverse reaction signal mining as Topics related to the top 10 in terms of country/region, institution, source of authorship, publication, annual text volume, annual citation volume, and quotation. At the same time, using the Chinese Journal Full-text Database in CNKI as a search platform, advanced search “adverse reaction signal mining” “adverse reaction data mining” as the subject of all Chinese literature and literature metrology analysis. The time span is from January 1st, 1990 to June 30th, 2018. **Results:** 191 English documents were retrieved from the Web of Science database. The number of annual and year citations was the highest in 2013, the h-index was 36, and 103 articles were published in the United States and France, accounting for 53.93% of the total. The number of Chinese documents retrieved in CNKI was 41, and the most studied was the Second Military Medical University. The number of documents issued was 11 and accounted for 26.83% of the total. **Conclusion:** The research on adverse drug reaction signal mining research increased rapidly, and has always maintained a certain degree of research interest, but the mining technology and data utilization needs to be further matured.

KEY WORDS Adverse drug reactions; Signal mining; Data mining; Bibliometrics

药品不良反应(adverse drug reaction, ADR)是药品在使用过程中导致的与用药目的无关或意外的有害反应,可以是预期的副作用,也可以是无法预期

的过敏性或特异性反应。据世界卫生组织(WHO)统计,因 ADR 住院的患者占住院人数 5% ~ 10%,而住院患者中发生 ADR 的人数达 10% ~ 20%,致

死率为 0.24% ~ 2.9%，ADR 日益成为一个严重的公共卫生问题^[1]。药品上市前评估选取患者的样本过小，观察时间有限，未考虑到食物及环境等因素的影响，很难发现一些迟发的、罕见的 ADR^[2]。因此，对 ADR 信号的发现和挖掘至关重要。ADR 信号指未知的或是尚未完全证明的药品与不良事件之间可能存在因果关系的报告信息^[3]。通过对 ADR 信号的大数据挖掘，可以获取药品与不良事件之间的潜在关联，为后续临床、药学、流行病学等多个学科的相关研究提供信息支持。对 ADR 信号的筛选和发生特点的探索是药物警戒和药品上市后安全性评价的重要研究内容^[4]。

文献计量分析有助于研究人员对相关研究领域的国内外情况有一个全面了解，包括研究能力较强的国家/地区、机构、科学家，载文量较多的杂志以及被引频次较多的文章等情况^[5]。进行文献计量学分析有助于科研选题、寻求合作伙伴、成果发表等，还可为相关专业学生的求学深造提供参考^[6]。本文以 ADR 信号挖掘相关的文献为研究对象，进行文献计量学研究，为相关人员提供参考。

1 资料与方法

1.1 数据库

以科学引文索引 (science citation index, SCI) 数据库 Web of Science 及中国知网 (CNKI) 中国期刊全文数据库为检索平台。

1.2 检索策略

①采用高级检索模式；②英文利用“adverse reaction signal mining”OR“adverse reaction data mining”为主题词进行高级检索，中文以“不良反应信号挖掘”“不良反应数据挖掘”为主题词检索；③所有相关文献，文献发表时间自建库截止至 2018 年 6 月 30 日。对检索结果通过阅读题目和摘要进行筛选，排除标准：①与“adverse reaction signal mining”“adverse reaction data mining”或“不良反应信号挖掘”“不良反应数据挖掘”无关的文章；②重复发表的文章；③非临床相关文章。

1.3 评价内容

对搜索获得的相关文献进行文献计量学分析，包括国家/地区、机构、作者、来源出版物、年出版文献量；对引文数、文献总被引、年均被引、h-index 以及被引频次排前 10 位的文献进行分析；并对中文文献进行了简要分析，如机构、来源出版物、引用频次排前 10 位的文章等。利用引文分析软件 (history of

cite, Hist Cite)^[7]对 ADR 信号挖掘研究领域文献进行文献计量分析和引文分析，生成引文编年图，展现该领域的发展历史、研究现状，帮助研究人员追踪最新的研究趋势，判断研究方向^[8]。

2 结果

在 Web of Science 数据库中检索到英文文献 243 篇，其中期刊论文 191 篇，其余为会议论文等。下载包含参考文献的全纪录题录到文本文档，对年发文章量、年引文数、引用总频次及发文章量排前 10 位的国家、机构和科学家进行分析。

2.1 年发文章量与年引文数

对检索到的 191 篇英文期刊论文进行汇总分析，年发文章量见图 1。从图中可以看出对 ADR 信号挖掘的研究始于 2009 年，此后数量逐渐增多，在 2013 年达到最高峰。图 2 为年引文量，在 2015 年最多。表 1 为被引频次与施引文献情况汇总。

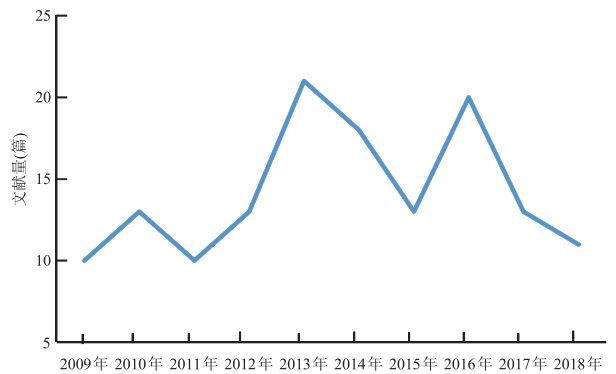


图 1 关于药品不良反应信号挖掘研究的年发文章量

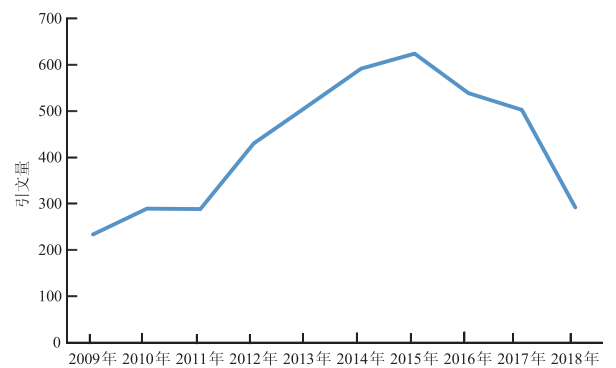


图 2 关于药品不良反应信号挖掘研究的年引文数

2.2 发文章量排前 10 位的国家/地区、机构、出版物和科学家

发文章数量的国家排名能客观反映相关国家/地区在该领域研究成果的数量情况。表 2 为发文章量排前 10 位的国家/地区排名，美国和法国为 ADR 信号

挖掘研究较多的国家,发文总量为 103 篇,占总数的 53.93%。

表 1 被引频次与施引文献情况汇总

项目	定义	记录数
找到的结果数		191
总被引频次	检索结果集中所有项目的总被引频次(引用参考文献),是所有年份施引文献总数全部相加后的总和	5220
去除自引的总被引频次	所有项目的被引频次总和(引用参考文献),减去由此集中的论文进行的任何引用	4656
施引文献	又叫来源文献,是引用引文的文献,即附有参考文献的原始文献	3269
去除自引的施引文献	施引文献总数,但不包括出现在“引文报告”检索结果集中的任何文献	3130
每项平均引用次数	“被引频次”计数总和除以查找到的结果数量	27.33
h-index	将论文依被引频次从高到低进行排序,h-index 第“h”个论文的施引文献数≥“h”	36

表 2 发文量排前 10 位的国家/地区

排名	国家/地区	记录数	比例(%)
1	美国	81	42.41
2	法国	22	11.52
3	瑞典	18	9.42
4	日本	17	8.90
5	新西兰	15	7.85
6	英国	14	7.33
7	澳大利亚	11	5.76
8	意大利	11	5.76
9	中国	8	4.19
10	西班牙	8	4.19

机构发文量排名可以了解到各个单位在该领域的研究成果情况。表 3 为发文量排前 10 位的机构以及所属国家/地区。从表中可以看出,排名并列第一的是美国食品药品监督管理局(FDA),发文量为 15 篇,占总数的 7.85%。

表 3 发文量排前 10 位的机构

排名	机构	所属国家/地区	文献数	比例(%)
1	美国食品药品监督管理局	美国	15	7.85
2	纽约大学	美国	12	6.28
3	辉瑞制药	美国	12	6.28
4	纽约医学院	美国	11	5.76
5	伊拉斯姆斯大学	荷兰	10	5.24
6	法国国家健康与医学研究院	法国	8	4.19
7	波尔多第一大学	法国	8	4.19
8	墨西拿大学	意大利	7	3.67
9	乌普萨拉监测中心	WHO	7	3.67
10	日本近畿大学	日本	6	3.14

表 4 为发文量排前 10 位的出版物排名,排名第一的为 Drug Safety。科学家排名及发文数量见表 5。

2.3 ADR 信号挖掘研究的引文编年图

文献集内被引用次数代表了在研究领域内被关注的程度。利用 Hist Cite 对文献集内被引用次数最

表 4 发文量排前 10 位的出版物

排名	出版物	文献数	比例(%)
1	Drug Safety	29	15.18
2	Pharmacoepidemiology and Drug Safety	11	5.76
3	European Journal of Clinical Pharmacology	10	5.24
4	Journal of Biomedical Informatics	7	3.67
5	Bmc Bioinformatics	6	3.14
6	Clinical Pharmacology Therapeutics	6	3.14
7	International Journal of Medical Sciences	6	3.14
8	Journal of the American Medical Informatics Association	6	3.14
9	Drug Information Journal	5	2.62
10	Vaccine	5	2.62

表 5 发文量排前 10 名的作者

排名	作者	文献数	比例(%)
1	Bate A	13	6.81
2	Hauben M	13	6.81
3	Lindquist M	9	4.71
4	Edwards Ir	8	4.19
5	He Hx	7	3.67
6	Kelman C	7	3.67
7	Reich L	7	3.67
8	Trifiro G	7	3.67
9	Cano M	6	3.14
10	Chen J	6	3.14

高的 10 篇文献生成可视化引文编年图。引文编年图共有节点 10 个,最小被引用频次为 13,最大被引用频次为 64。图中圆圈内的数字表示当前节点在文献集内的序号,圆圈的大小代表文献集内被引用次数的高低,箭头指向的文献是被引用的文献,左侧数字表示发表年份。见图 3。

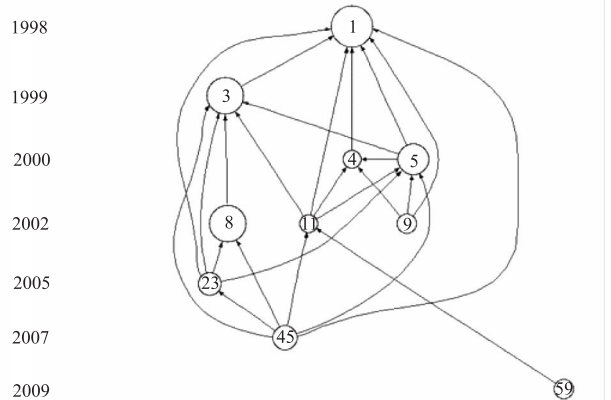


图 3 药品不良反应信号挖掘研究的引文编年图

从图中可以看出,1 号文献,即 Bate 等^[9]撰写的《贝叶斯神经网络方法:一种用于药物不良反应信号挖掘的方法》,被引用的次数最高。WHO 药物不良反应数据库包含近 200 万份报告,是世界上最大的此类数据库,每季度新增约 35 000 份报告。该文献介绍了贝叶斯神经网络方法是一种用于 ADR 信号挖掘的方法,是一种灵活的、自动的程序,可用于检测来自 WHO 国际药物监测中心数据集的重要

信号,找到具有已知概率差异的新信号。

居于被引用频次次位的是 8 号文献,由 Szarfman 等^[10]撰写的《使用筛选算法和计算机系统在 FDA 自发性报告数据库中预测药物和不良事件组合》,该文献通过对 FDA MedWatch 数据库中 ADR 自愿报告的多项目的 Gamma Poisson Shrinker (MGPS) 程序进行介绍,该程序提供了一个客观的、系统的数据视图,提醒重要的、新的安全信号,证实了数据挖掘在早期安全信号检测中的重要性及特异性。

2.4 中文文献的计量分析

对检索得到的有关 ADR 信号为挖掘的中文文献进行汇总分析,共有 41 篇文献,年发文量如图 4 所示,2016 年的发文量最高。第二军医大学的发文量最多,共有 11 篇。表 6 和表 7 分别为发文量的机构排名和作者排名。总被引次数排前 10 位的中文文献见表 8。

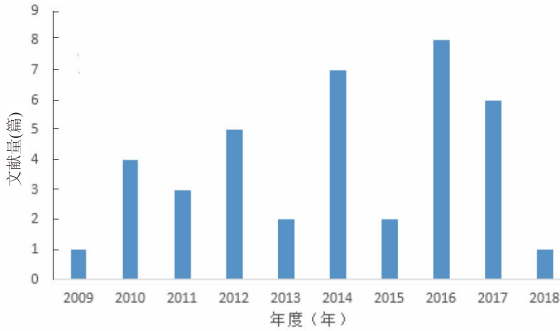


图 4 中文发表不良反应信号挖掘研究的年发文量情况

表 6 发文量排前 10 位的中国机构排名

排名	机构	文献数	占总数的比例 (%)
1	第二军医大学	11	26.83
2	北京中医药大学	5	12.19
3	天津大学	4	9.76
4	绵阳市中心医院	3	7.32
5	吉林大学	2	4.88
6	四川大学华西医院	2	4.88
7	山东大学	1	2.44
8	南京中医药大学	1	2.44
9	中国医学科学院医学科学院	1	2.44
10	山东省淄博市妇幼保健院	1	2.44

表 7 中文文献发表量前 10 名的作者排名

排名	作者	文献数	占总数的比例 (%)
1	郭军	3	7.32
2	王宇	3	7.32
3	叶小飞	3	7.32
4	贺佳	2	4.88
5	吴嘉瑞	2	4.88
6	蒋朋利	2	4.88
7	饶友义	2	4.88
8	潘建玲	2	4.88
9	储文功	2	4.88
10	李文军	1	2.44

表 8 总被引次数排名前 10 位的中文文献 (含学位论文)

文章题目	作者	出版单位名称	出版年	合计引用次数
基于文献数据库和传统药物警戒思想的中药注射剂安全性研究	吴嘉瑞	北京中医药大学	2007	31
药品不良反应信号检测方法理论及应用研究	李婵娟	第四军医大学	2008	26
检测联合用药不良反应信号的数据挖掘方法	钱铁峰	中国卫生统计	2010	14
基于数据挖掘的中药安全性研究与思考	童元元	中国中药杂志	2011	13
基于自发呈报系统与循证医学的药品不良反应信号挖掘	叶小飞	第二军医大学	2011	10
基于信息学理念的中药注射剂不良反应发生规律研究思路探讨	吴嘉瑞	中国执业药师	2009	9
基于数据挖掘技术的心脑血管用药 ADR 关联模型构建研究	冯变玲	天津大学	2012	8
他汀类药物的使用与认知能力下降、糖尿病及癌症之间的关系	蒋朋利	中国临床药理学与治疗学	2014	7
药品不良反应信号的验证和统计分析研究方法的探索——基于克林霉素注射剂致肾毒性的不良反应监测数据	路长飞	中国药物警戒	2013	7
数据挖掘技术在药品疗效上的应用	魏鲁霞	医学信息学杂志	2010	4

3 讨论

ADR 严重影响着广大患者的用药安全,对 ADR 信号的挖掘和检测是发现 ADR 的关键。可通过对上市后药品的监测,尤其是对 ADR 数据的分析,发现 ADR 信号可弥补药物临床研究的不足^[11]。

如何科学有效地分析和利用 ADR 数据库资料,建立快速有效的数据挖掘技术为药品风险管理及评价提供有价值的信号,及时、快速、准确地发现危害人类健康的危险药物,最大程度地减少其对人类的伤害,成为目前亟待解决的问题。运用数据挖掘技术从海量数据中提取出 ADR 信号,同时利用循证医学的方法科学地加以评价,已成为 ADR 研究方向的重点和热点。

目前研究的关于 ADR 信号挖掘英文文献,所有年份施引文献总数全部相加后的总和为 5 220。去除自引的总被引频次 4 656,指所有项目的被引频次总和,减去由此集中的论文进行的任何引用。该指标可以客观地说明该类学科研究论文总体被使用和受重视的程度,以及在学术交流中的作用和地位。施引文献 3 269,是引用引文的文献,即附有参考文献的原始文献。施引文献是被引文献的发展,可以大致称为研究的前沿。h-index 指将论文依被引频

次从高到低进行排序,是一种评价学术成就的新方法,h代表“高引用次数”(high citations),h-index越高,表明论文影响力越大。该领域的文章每篇被引用了至少36次的论文总共有36篇。引文编年图共有节点10个,文献集内被引用次数代表了在研究领域内被关注的程度。对引用频次较多的文献进行分析,发现目前关于ADR信号挖掘的研究主要集中在利用自发呈报系统进行ADR信号检测以及利用贝叶斯法进行数据挖掘。

近10年来,ADR信号挖掘研究发展迅速,但我国在该领域发表的英文文献仅有8篇,占全部文献数量的4.19%,中文相关文献41篇,可见我国在ADR信号挖掘研究上较美国等其他发达国家还有差距,在ADR信号挖掘方法方面应进一步深入探讨。本文利用文献计量学与Hist Cite方法能够快速、直观地揭示ADR信号挖掘研究领域中经典的研究文献和重要的研究发现,展现ADR信号挖掘研究的发展轨迹,同时也为进一步利用信号挖掘的方法进行ADR研究提供参考依据。

参 考 文 献

- 1 叶小飞. 基于自发呈报系统与循证医学的药品不良反应信号挖掘[D]. 上海:第二军医大学博士学位论文,2011
- 2 彭媛,王程程,唐利,等. 西格列汀上市后致皮肤不良反应的信号挖掘与评价[J]. 中国药房,2014,25(2):156-158
- 3 李婵娟. 药品不良反应信号检测方法理论及应用研究[D]. 西安:第四军医大学博士学位论文,2008
- 4 王宇,饶友义,郭军,等. 报告比值比法挖掘富马酸替诺福韦二吡呋酯不良反应信号[J]. 中国药房,2016,27(32):4515-4518
- 5 陈冠初. 文献计量学与非文献计量学在期刊评价中的应用[J]. 编辑学报,2006,18(6):472
- 6 丁月霞,李真,任瑞珍,等. 利拉鲁肽的文献计量学分析[J]. 药物流行病学杂志,2016,25(11):731-736
- 7 Garfield E, Paris S, Stock WG. A software tool for informetric analysis of citation linkage[J]. Information-Wissenschaft und Praxis, 2006, 57(8):391-400
- 8 宫建,李翠英,毕开顺,等. 药物流行病学在临床药学研究领域的可视化引文分析[J]. 药物流行病学杂志,2015,24(5):302-305
- 9 Bate A, Lindquist M, Edwards IR, et al. A Bayesian neural network method for adverse drug reaction signal generation[J]. Eur J Clin Pharmacol, 1998, 54(4):315-321
- 10 Szarfman A, Machado SG, O'Neill RT. Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the US FDA's spontaneous reports database[J]. Drug Safety, 2002, 25(6):381-392
- 11 张晓兰,夏佳. 浅谈药物警戒中的安全信号与信号管理[J]. 药物流行病学杂志,2012,21(2):90-94