

· 数据挖掘 ·

基于数据挖掘的癫痫用药规律探索

武凤震, 赵文光*

(广州中医药大学医学信息工程学院, 广州 510006)

[摘要] **目的:**挖掘《中医方剂大辞典》中治疗癫痫方剂的用药规律。**方法:**收集、整理《中医方剂大辞典》中收录的治疗癫痫方剂,建立癫痫方剂文献研究数据库,采用数据挖掘技术(频数分析、聚类分析、关联规则)对其用药规律进行分析。**结果:**关于治疗癫痫的文献中,共涉及532首方剂和488味中药,治疗癫痫的核心药物有朱砂、甘草、人参、牛黄等,用药种类以补虚药、安神药、平肝熄风药、清热药、化痰药为主。对核心药物聚类分析,形成了6个聚类组合。通过关联规则分析,共得22条药对及药组关联规则。**结论:**癫痫方剂的用药规律以祛风止痉、疏肝行气、开窍醒神、活血化痰、养心安神为主,中医治疗癫痫应针对肝风、郁火、积痰、血瘀、气乱等因素进行辨证论治,同时还应考虑正气亏虚的因素。

[关键词] 癫痫;《中医方剂大辞典》;数据挖掘;用药规律

[中图分类号] R287 **[文献标识码]** A **[文章编号]** 1005-9903(2017)08-0190-06

[doi] 10.13422/j.cnki.syfx.2017080190

[网络出版地址] <http://www.cnki.net/kcms/detail/11.3495.R.20170208.1942.022.html>

[网络出版时间] 2017-02-08 19:42

Analysis on Medication Regularity of Epilepsy Medicine Based on Data Mining Technology

WU Feng-zhen, ZHAO Wen-guang*

(College of Medical Information Engineering, Guangzhou University of
Chinese Medicine, Guangzhou 510006, China)

[Abstract] **Objective:** To summarize the medication principles of prescriptions for epilepsy from *Prescription Dictionary of Traditional Chinese Medicine (TCM)*. **Method:** Prescriptions used for epilepsy were collected from the *Prescription Dictionary of Traditional Chinese Medicine* to establish the literature research database of Chinese medicine prescriptions in treating epilepsy and analyze the medication principles of prescriptions through data mining (frequency analysis, clustering analysis, correlation analysis). **Result:** A total of 532 prescriptions and 488 herbs were involved in the literatures about treating epilepsy. The herbs with high usage frequency in the prescriptions included Cinnabaris, Glycyrrhizae Radix et Rhizoma, Ginseng Radix Et Rhizoma, Bovis Calculus; and deficiency-nourishing drugs, sedative drugs, liver wind calming drugs, clearing-heat drugs and expectorant drugs were mainly used. Six clustering combinations can be formed based on clustering analysis of high frequency drugs. Through correlation analysis, we obtained 22 correlated rules of medicine pairs and combinations. **Conclusion:** The prescription regularity of epilepsy medicines is dominated by wind-dispelling and tetany-relieving, live-coursing and Qi moving, orifices-opening and spirit-arousing, blood-quickening and stasis-transforming, and tranquilizing mind by nourishing the heart. TCM for treating epilepsy shall be based on syndrome differentiation caused by such factors as liver wind, stagnated fire, accumulation of phlegm, blood stasis and Qi-movement

[收稿日期] 20161030(005)

[基金项目] 广东省省级科技计划项目(2014A020221086);广东省中医药科学院联合科研项目(2014A020221039)

[第一作者] 武凤震,在读硕士,从事中医药数据挖掘研究,Tel:13226651902,E-mail:26006968080@qq.com

[通讯作者] *赵文光,硕士,副教授,从事中医药数据挖掘研究,Tel:020-39358151,E-mail:1247112316@qq.com

disturbance, and give consideration to the factor of the deficiency of righteousness.

[**Key words**] epilepsy; *Prescription Dictionary of Traditional Chinese Medicine*; data mining; medication regularity

癫痫是阵发性、暂时性脑功能失调,是一种神经内科常见病。主要是由于大脑神经元反复、过度的同步化放电并导致发作性、突然性、短暂性的脑功能紊乱^[1]。国内近年来的临床流行病学研究显示,癫痫的患病率在 0.5% 左右^[2]。癫痫具有病因复杂,发病机制尚不明确,病程长,难根治等特点,是中医常见的疑难病证之一,亦被西医列为现代难治病^[3]。目前西医对癫痫的治疗方法主要有药物治疗、手术治疗、神经调控治疗,存在药物副作用大、术后复发率高等问题^[4]。根据其发作临床表现和特征,癫痫属于中医的“痫病”“痫证”等范畴。中医治疗癫痫具有悠久的历史和丰富的经验,同时也积累了大量的文献资料。其中方剂是中医药临床应用的主要形式和手段,对方剂的用药规律进行总结,是传承中医防治疾病方法的关键^[5]。

数据挖掘^[6]是从大量的、不完全的、有噪声的、模糊的、随机的实际应用数据中,提取隐含在其中的、未知的、潜在有用的信息和知识的过程^[7]。数据挖掘技术可以从海量数据中提取和发现规律性的认识和隐含的知识^[8]。在中医药领域,随着人们健康、疾病管理数据及传统医药数据的电子化、结构化,数据挖掘技术的应用日益增多^[9]。数据挖掘在中医药领域的研究遍及特定疾病、方剂、医案、名老专家经验、中药以及中医诊断学等多个方面,其中特定疾病、方剂、医案以及名老专家经验是其中的热点^[10]。例如何佳等^[11]收集中药复方防治脂肪肝的文献,建立数据库,利用数据挖掘技术总结了中药复方防治脂肪肝的用药规律;唐仕欢等^[12]应用中医传承辅助系统软件,使用关联规则、改进的互信息法、复杂系统熵聚类等数据挖掘方法,分析了哮喘方剂中常用的药物、组合规则、核心组合及发现了新方。李园白等^[13]对 664 例医案进行关联规则分析,总结了崩漏症状、证型、用药之间的关系。王珊珊等^[14]运用频数分析与关联规则分析,从症状、病理因素、病机、处方用药等多方面对周仲瑛教授治疗肺癌的医案进行数据挖掘。故本研究通过收集《中医方剂大辞典》^[15]中治疗癫痫的方剂数据,建立癫痫方剂数据库并对其进行数据挖掘,总结文献中治疗癫痫的用药规律进而为该病的临床治疗提供参考。

1 资料与方法

1.1 数据来源与规范化 《中医方剂大辞典》是将自秦汉到现代的中医药著作中的方剂进行整理、研究、编纂而成的方剂学大型工具书。全书约 1 800 万字,载方约 10 万首。该书具有数据量大,数据规范性强的特点,故本研究以《中医方剂大辞典》作为治疗癫痫方剂的数据来源。通过查询主题词“癫痫”,收集整理后共得有效方剂 532 首。建立癫痫方剂数据库,录入包括方剂编号、方名、方源、剂型、主治证型及药物组成在内的方剂信息。以《中国药典》^[16]2015 年版一部以及《中药大辞典》^[17]中记载的中药正名为标准规范方剂中药物名称。例如把光明砂、辰砂、丹砂等规范为朱砂,把生军、川军、锦纹大黄等规范为大黄。如果出现多个中药名称合并的现象,将其拆分并规范为中药正名,如将“焦三仙”拆分并规范为山楂、神曲、麦芽,将“朱茯神”拆分并规范为朱砂、茯神。

1.2 资料处理与数据分析 通过 Excel 2013 建立癫痫方剂数据库,并将数据转化为计算机能识别的格式,在此基础上对所有的药物进行频数统计。采用 IBM SPSS Statistics 22 软件中的 Hierarchical 聚类算法对核心药物进行聚类分析,以及通过 IBM SPSS Modeler 17 数据挖掘软件进行建模,采用软件内置的 Apriori 算法对核心药物进行关联规则分析。

2 结果

2.1 药物频数、频率分析 纳入分析的 532 首方剂中涉及中药共 488 味,药物出现频数总计 4 956 次,使用频率最高的是朱砂,共使用 184 次,频率为 34.59%;其次为甘草,其使用频数为 142,使用频率为 26.69%;人参位居第 3 位,其使用频数为 122,使用频率为 22.93%;牛黄共使用 111 次,其使用频率为 20.86%;位居第 4 位;麝香使用频数为 104,使用频率为 19.55%,位居第 5 位,以上 5 味药使用频数均超过 100。其中,使用频数排列在前 10 位的中药分别是朱砂、甘草、人参、牛黄、麝香、远志、大黄、茯苓、半夏、黄芩,其中朱砂、甘草、人参、牛黄 4 味药物使用频率超过 20%。使用频率 > 20%,使用频数 > 40 次的药物见表 1。

2.2 药类使用频数、频率分析 根据《中国药典》2015 年版与《中药大辞典》中药物分类标准,将频数

表 1 使用频数 >40 次的药物

Table 1 Herbs used for over 40 times in prescriptions for epilepsy

No.	药物	频数/次	频率/%	No.	药物	频数/次	频率/%
1	朱砂	184	34.59	20	全蝎	59	11.09
2	甘草	142	26.69	21	僵蚕	56	10.52
3	人参	122	22.93	22	犀角	56	10.52
4	牛黄	111	20.86	23	麦冬	54	10.15
5	麝香	104	19.55	24	白术	53	9.96
6	远志	89	16.73	25	胆南星	53	9.77
7	大黄	88	16.54	26	石膏	52	9.21
8	茯苓	88	16.54	27	白芍	49	9.21
9	半夏	87	16.35	28	龙齿	49	9.21
10	黄芩	85	15.98	29	肉桂	49	9.21
11	附子	84	15.80	30	陈皮	47	8.83
12	防风	79	14.85	31	酸枣仁	46	8.65
13	当归	78	14.66	32	细辛	46	8.65
14	茯神	78	14.66	33	川芎	45	8.46
15	天南星	76	14.29	34	钩藤	45	8.46
16	黄连	74	13.91	35	地黄	44	8.27
17	石菖蒲	73	13.72	36	冰片	42	7.89
18	天麻	70	13.16	37	金箔	42	7.89
19	雄黄	63	11.84	38	麻黄	42	7.89

表 2 每类药物使用频数和频率分布

Table 2 Use frequency and frequency distribution of each type of herb

No.	药物类别(味数)	频数/次	频率/%	累积频率/%
1	补虚药(10)	548	11.06	11.06
2	安神药(12)	546	11.02	22.08
3	平肝熄风药(11)	487	9.83	31.91
4	清热药(13)	429	8.66	40.57
5	化痰药(11)	380	7.67	48.24
6	解表药(11)	320	6.46	54.70
7	开窍药(3)	219	4.42	59.12
8	利水渗湿药(3)	183	3.69	62.81
9	理气药(8)	180	3.63	66.44
10	温里药(4)	180	3.63	70.07
11	攻下药(4)	147	2.97	73.04
12	攻毒杀虫止痒药(3)	119	2.40	75.44
13	祛风湿药(6)	117	2.36	77.80
14	活血化瘀药(4)	105	2.12	79.92
15	拔毒化腐生肌药(2)	42	0.81	80.73
16	驱虫药(1)	14	0.28	81.01
17	收涩药(1)	11	0.22	81.23
18	化湿药(1)	10	0.20	81.43

10 次以上的 108 味高频药物进行分类。具体药物类别、频数、频率分布见表 2。由表 2 可见,治疗癫痫方剂中药物类别主要为补虚药、安神药、平肝熄风药、清热药、化痰药、解表药、开窍药、利水渗湿药、理气药、温里药,累计频率可达 70.07%。其中补虚药以使用 548 次,占总使用次数的 11.06% 高居第 1 位;安神药使用频数仅次于补虚药为 546 次,占总使用频率的 11.02%,位居第 2 位;平肝熄风药使用频数为 487 次,占总使用频率的 9.83%,位居第 3 位;清热药使用频数为 429 次,占总使用频率的 8.66%,位居第 4 位;化痰药使用 380 次,占总使用次数的 7.67%,位居第 5 位;解表药使用频数为 320 次,占总使用次数的 6.46%,位居第 6 位;开窍药使用频数为 219 次,占总使用次数的 4.42%,位居第 7 位;利水渗湿药使用频数为 183 次,占总使用次数的 3.69%,位居第 8 位;理气药与温里药使用频数都为 180 次,占总使用次数的 3.63%,并列排名第 9。

2.3 聚类分析结果 采用 IBM SPSS Statistics 22 软件中的 Hierarchical 聚类算法对上述 38 味核心药物进行聚类分析,采用系统聚类将其聚为 2~10 类,结合相关中医理论,经过分析后认为聚成 6 类比较

合适。这 6 类分别为聚一类:朱砂、牛黄、麝香、雄黄、犀角、麦冬、金箔、冰片;聚二类:甘草、人参、远志、茯苓、茯神、当归、黄连、石菖蒲、白术、白芍、酸枣仁、陈皮、地黄;聚三类:大黄、黄芩、石膏、龙齿、钩藤、麻黄;聚四类:半夏、附子、天南星、天麻、全蝎、僵蚕;聚五类:防风、肉桂、细辛、川芎;聚六类:胆南星。药物聚类分布见图 1。

2.4 关联规则分析结果 采用 IBM SPSS Modeler 17 数据挖掘软件进行建模,运用软件内置的 Apriori 算法运对上述核心药物进行关联规则分析,设置参数:支持度 $\geq 7\%$,置信度 $\geq 60\%$ (支持度表示关联规则的普遍性,即数据库中 $X \cup Y$ 的事务占全部事务的百分比;置信度是对关联规则准确度的测量,表示包含 $X \cup Y$ 的事务数与包含 X 的事务数的比值^[18])。共得出 22 条药对与药组关联规则,关联规则见图 2,表 3。由表 3 可见,朱砂->牛黄支持度最高,其支持度为 20.64%;远志->茯神,人参置信度最高,其置信度为 79.07%;朱砂->牛黄规则支持最高,其规则支持为 12.38%;茯神->远志,人参提升度最高,其提升度为 5.01;朱砂->牛黄部署能力最高,其部署能力为 8.26;朱砂->麝香,牛黄规则 ID

表 3 癫痫方剂数据关联规则分析

Table 3 Association rules analysis of data of epilepsy prescriptions

No.	关联规则	支持度/%	置信度/%	规则支持/%	提升度	部署能力	规则 ID	关联数/次
1	朱砂->牛黄	20.64	60.00	12.38	1.82	8.26	17	110
2	朱砂->麝香	19.33	63.11	12.20	1.91	7.13	16	103
3	远志->茯神	14.45	97.53	9.76	4.09	4.69	14	77
4	甘草->当归	14.45	61.04	8.82	2.34	5.63	15	77
5	朱砂->雄黄	11.63	69.36	8.07	2.10	3.57	10	62
6	牛黄->犀角	10.32	60.00	6.19	2.91	4.13	13	55
7	附子->僵蚕	10.13	64.82	6.57	4.37	3.57	12	54
8	朱砂->麝香,牛黄	9.94	69.82	6.94	2.11	4.13	22	53
9	甘草->白术	9.76	65.39	6.38	2.51	3.57	11	52
10	人参->茯神,远志	9.76	65.39	6.38	2.95	3.00	19	52
11	茯神->远志,人参	8.82	72.34	6.38	5.01	3.38	21	47
12	当归->白芍	8.44	62.22	5.25	4.31	3.38	4	45
13	甘草->白芍	8.44	66.67	5.63	2.56	2.44	5	45
14	远志->酸枣仁	8.44	64.45	5.44	3.90	3.19	6	45
15	当归->川芎	8.26	61.36	5.07	4.25	2.81	9	44
16	麝香->雄黄,朱砂	8.07	62.79	5.07	3.25	3.00	18	43
17	远志->茯神,人参	8.07	79.07	6.38	4.79	1.69	20	43
18	牛黄->金箔	7.88	61.91	4.88	3.00	3.00	7	42
19	朱砂->金箔	7.88	69.05	5.44	2.09	2.44	8	42
20	麝香->冰片	7.51	75.00	5.63	3.88	1.88	1	40
21	牛黄->冰片	7.51	65.00	4.88	3.15	2.63	2	40
22	朱砂->冰片	7.51	72.50	5.44	2.20	2.06	3	40

酸枣仁;清热药以清热燥湿药、清热泻火药、清热凉血药为主,代表药为黄芩、黄连、石膏、地黄;开窍药中麝香使用频率最高;化痰药主要以温化寒痰药为主,代表药为半夏、天南星。

3.2 药物配伍规律分析

3.2.1 聚类分析 聚类分析是依据数据的某些属性将其聚集成类,使类间相似性尽量小,类内相似性尽量大。聚类分析是一种无监督的学习,无需任何先验知识,直接从数据对象中发现有意义的结构,进而把输入对象被划分为一个未知的类。聚类分析能够作为一个独立的工具获得数据的分布状况^[24]。对中医治疗癫痫方剂中的核心药物进行聚类分析形成 6 个聚类方,分析可见,聚一类与聚三类主要是重镇安神药、平肝熄风药、开窍醒脑药、清热补虚药相配伍,主要针对肝风、郁火这两个病理因素进行治疗。聚二类以补虚药、养心安神药、理气药、开窍药、清热药、利水渗湿药相配伍,主要针对气乱、郁火这两个病理因素进行治疗。聚四类中半夏与天南星温

化寒痰,天麻、全蝎与僵蚕平肝潜阳、息风止痉,主要针对痰火扰神、肝阳上亢型癫痫。聚五类与聚六类功效分别以活血化瘀、温化寒痰为主。由此可见,癫痫方剂的组方以祛风止痉、疏肝行气、开窍醒神、活血化瘀、养心安神为主,除此之外,聚类分析还显示应注重补虚药的使用。

3.2.2 关联规则分析 关联规则又称关联挖掘,是从大量的数据中发现项集之间的频繁模式、相关性或因果结构,从而描述一个事物中某些属性同时出现的规律和模式的分析技术^[25]。对中医治疗癫痫方剂中的核心药物进行关联规则分析,分析结果可见,支持度较高的药物组合有朱砂->牛黄;朱砂->麝香;远志->茯神;甘草->当归;朱砂->雄黄;牛黄->犀角;附子->僵蚕;朱砂->麝香,牛黄;甘草->白术;人参->茯神,远志。这些药对与药组主要是安神药、开窍药、补虚药、平肝熄风药、清热药的不同组合。如安神药->平肝熄风药(朱砂->牛黄),安神药->开窍药(朱砂->麝香),补虚药->补

虚药(甘草->当归)、平肝熄风药->清热药(牛黄->犀角),安神药->开窍药->平肝熄风药(朱砂->麝香->牛黄)。上述结果体现了治疗癫痫的重镇安神、平肝熄风、开窍醒神、补虚安神、清热泻火等治则治法。

综上所述,治疗癫痫方剂的数据挖掘结果与癫痫的病因病机及治则治法基本相符,中医治疗癫痫应针对肝风、郁火、积痰、血瘀、气乱等因素进行辨证论治,同时应考虑患者正气亏虚的因素,进而达到治疗癫痫的目的。

[参考文献]

[1] 王拥军,丁成云. 建立中国癫痫临床诊疗指南的意义[J]. 北京医学,2005,27(11):678-680.

[2] 常琳,王小珊. 中国癫痫流行病学调查研究进展[J]. 国际神经病学神经外科学杂志,2012,39(2):161-164.

[3] 黄斌. 癫痫中西医论治[M]. 北京:人民卫生出版社,1999:3.

[4] 司富春,宋雪杰,李洁,等. 癫痫证候和方药分布规律文献分析[J]. 中医杂志,2014,55(6):508-512.

[5] 唐仕欢,杨洪军. 中医组方用药规律研究进展述评[J]. 中国实验方剂学杂志,2013,19(5):359-363.

[6] HAN J W, Kamber M, PEI J. 数据挖掘概念与技术[M]. 北京:机械工业出版社,2012:4-5.

[7] 王国平. IBM SPSS Modeler 数据与文本挖掘实战[M]. 北京:清华大学出版社,2014:4.

[8] 刘树春,刘洋,张晓玮,等. 基于方剂数据的补肾常用中药及其配伍规律的挖掘分析[J]. 中国实验方剂学杂志,2015,21(20):208-212.

[9] 朱彦,朱玲,崔蒙. 论中医药数据挖掘[J]. 中华中医药杂志,2016,20(8):2932-2935.

[10] 李强. 基于方剂数据挖掘的痹证证治规律研究[D]. 北京:中国中医科学院,2014.

[11] 何佳,卢秉久. 基于现代文献的中药复方防治脂肪肝

用药规律探究[J]. 辽宁中医杂志,2016,50(7):1364-1367.

[12] 唐仕欢,卢朋,李健,等.《中医方剂大辞典》中治疗哮喘病方剂组方规律分析[J]. 中医杂志,2012,53(14):1230-1233.

[13] 李园白,杨阳,崔蒙,等. 664例崩漏医案信息的数据挖掘[J]. 中国中医药信息杂志,2009,16(11):95-96.

[14] 王珊珊,郭茗,朱垚,等. 基于数据挖掘探讨周仲瑛对肺癌的辨治思路[J]. 辽宁中医杂志,2016,50(8):1581-1584.

[15] 彭怀仁. 中医方剂大辞典[M]. 北京:人民卫生出版社,1994:1-12885.

[16] 国家药典委员会. 中华人民共和国药典. 一部[M]. 北京:中国医药科技出版社,2015:1-1268.

[17] 南京中医药大学. 中药大辞典[M]. 上海:上海科学技术出版社,2006:1-3875.

[18] 薛薇,陈欢歌. SPSS Modeler 数据挖掘方法及应用[M]. 北京:电子工业出版社,2014:209.

[19] 王永炎,张伯礼. 中医脑病学[M]. 北京:人民卫生出版社,2007:774.

[20] 雪瑛,崔树森. 癫痫病因病机探讨[J]. 河南中医,2000,20(5):16-17.

[21] 孙守华. 癫痫病机证治探要[J]. 甘肃中医,1999(6):3-5.

[22] 刘星,王欢. 癫痫发作期本虚标实病机探讨[J]. 中医药研究,1999(6):8-9.

[23] 王永炎,鲁兆麟. 中医内科学[M]. 北京:人民卫生出版社,2011:321.

[24] 刘兴方,韩学杰. 数据挖掘技术在医案研究中的应用与讨论[J]. 中国实验方剂学杂志,2014,20(9):247-250.

[25] 熊平. 数据挖掘算法与Clementine实践[M]. 北京:清华大学出版社,2011:75-77.

[责任编辑 邹晓翠]