

· 不良反应 ·

8 种胃癌化疗药物与药品不良反应的聚类关系分析

郭佳栋^{1,2}, 张雪梅^{1,2}, 刘影^{1,2}, 冯变玲^{1,2*} (1. 西安交通大学药学院药事管理与临床药理学系, 西安 710061; 2. 西安交通大学药品安全与政策研究中心, 西安 710061)

摘要 目的: 分析胃癌化疗药物不良反应与药品品种间的聚类关系, 促进临床安全合理用药。**方法:** 收集 109 家药品生产企业 2010-2014 年生产的 8 种胃癌化疗药物氟尿嘧啶、卡培他滨、表柔比星、顺铂、伊立替康、紫杉醇、奥沙利铂和多西他赛的不良反报告数据, 运用数据挖掘技术对药品不良反应与药品品种间的关系进行聚类分析, 使用 Clementine 14.1 软件进行统计。**结果与结论:** 共收集有效报告 10572 份, 卡培他滨和顺铂, 表柔比星和伊立替康, 多西他赛和紫杉醇联合用药发生一般药品不良反应的可能性较大。卡培他滨、顺铂、伊立替康和表柔比星联合用药; 多西他赛、氟尿嘧啶和紫杉醇联合用药时, 出现严重药品不良反应的概率较高。瘙痒、过敏性休克、顺铂、多西他赛在聚类过程中的运算比重较大。

关键词: 药品不良反应; 胃癌化学药物; 氟尿嘧啶; 卡培他滨; 表柔比星; 顺铂; 伊立替康; 紫杉醇; 奥沙利铂; 多西他赛; 聚类关系; 数据挖掘

中图分类号: R256.35; R725.7; R994.11 文献标识码: A 文章编号: 1002-7777(2017)05-0584-07
doi:10.16153/j.1002-7777.2017.05.022

Cluster Analysis of Eight Kinds of Gastric Cancer Drugs and Adverse Drug Reactions

Guo Jiadong^{1,2}, Zhang Xuemei^{1,2}, Liu Ying^{1,2}, Feng Bianling^{1,2*} (1. The Department of Pharmacy Administration and Clinical Pharmacy, School of Pharmacy, Xi'an Jiaotong University, Xi'an 710061, China; 2. The Center for Drug Safety and Policy Research, Xi'an Jiaotong University, Xi'an 710061, China)

Abstract Objective: To find the potential cluster relationship between drugs and adverse drug reactions, so as to build the cluster model and guide the clinical drug use. **Methods:** Reports of adverse reactions caused by 8 kinds of gastric cancer drugs were collected from 109 pharmaceutical producing enterprises during 2010-2014, including fluorouracil, capecitabine, epirubicin, irinotecan, cisplatin, paclitaxel, oxaliplatin and docetaxel. Data mining technology was used to build cluster model and clementine 14.1 was for statistical analysis. **Results and conclusion:** 10,572 effective reports were collected and statistically analyzed, which showed that capecitabine combined with cisplatin, epirubicin with irinotecan, and docetaxel with paclitaxel were more likely to cause general adverse drug reactions. The combination of capecitabine, cisplatin, irinotecan and epirubicin, as well as the combination of docetaxel, fluorouracil and paclitaxel, was more likely to cause serious adverse drug reactions. Pruritus, anaphylactic, cisplatin and docetaxel were more important in the process of cluster analysis.

Keywords: adverse drug reactions; gastric cancer chemotherapy drug; fluorouracil; capecitabine; epirubicin; irinotecan; cisplatin; paclitaxel; oxaliplatin; docetaxel; cluster analysis; data mining

基金项目: 国家社科基金资助项目(编号 13BGL132); 陕西省社会发展科技攻关项目(编号 2016SF-159)

作者简介: 郭佳栋, 硕士研究生; 研究方向: 药事管理; E-mail: gjdxjtu@163.com

通信作者: 冯变玲, 副教授; 研究方向: 药事管理; E-mail: fengbl@mail.xjtu.edu.cn

胃癌是一种发生在胃上皮组织的恶性肿瘤，是目前最常见的恶性肿瘤之一^[1]。近年来新发胃癌患者更加年轻化，30岁以下胃癌患者由1970年的1.7%上升至如今的3.3%，这一年轻化的趋势严重影响了社会生产的顺利进行，给家庭和社会带来了巨大负担^[2]。药品不良反应（adverse drug reaction, ADR）是指合格药品在正常用法用量下出现的与用药目的无关的或意外的有害反应。由于ADR可能对人体健康产生不可逆的损害，对于ADR的监测和ADR制度的完善越来越受到全球各个国家的高度重视。

数据挖掘是一个利用各种方法，从海量数据中提取对决策有用信息的过程^[3]。数据挖掘中的K-Means聚类也称快速聚类，属于覆盖型数值划分聚类算法^[4]。本研究运用K-Means聚类，对相应药品及ADR进行聚类分析，以期望寻找到药品与不良反应间的聚类特点，为促进临床合理用药及新药研发提供参考。

1 资料与方法

1.1 数据来源

本研究以2010-2014年收集到的109家药品生产企业生产的氟尿嘧啶、卡培他滨、表柔比星、顺铂、伊立替康、紫杉醇、奥沙利铂和多西他赛化疗药物在治疗胃癌过程中的ADR报告为研究对

象，运用数据挖掘技术中K-Means聚类算法，使用Clementine 14.1软件进行统计分析。

1.2 数据清洗

数据清洗的目的在于提升数据质量，使之符合所选择分析技术的要求。为此，在数据挖掘算法执行之前，必须对收集到的原始数据进行预处理，摒弃一些与挖掘目标不相关的属性，为数据挖掘提供干净、准确、更有针对性的数据，提高挖掘效率。

2 结果

2.1 ADR报告筛选

对于获得的数据，首先剔除非胃癌疾病患者，然后剔除重复报告，最后剔除年龄、性别等信息不详的数据，同时参照《WHO药品不良反应术语集》中对ADR术语的规定，将不规范的不良反应术语进行标准化处理。最后确定的有效报告为10572份，详见表1。

2.2 按频次挖掘的K-Means聚类结果

2.2.1 支持度筛选

支持度表示一组数据集在统计意义上需要满足的最低程度。为了研究药品在使用人群中发生不良反应的普遍性，按频率将原患疾病及不良反应排序并选取支持度 ≥ 0.01 ，共计4206份病例报告符合支持度的要求，详见表2。

表 1 8种抗肿瘤药物ADR报告分布

药品种类	清洗后数据	
	频数 / 份	比例 / %
奥沙利铂	4741	44.8
表柔比星	201	1.9
多西他赛	871	8.2
氟尿嘧啶	2018	19.1
卡培他滨	650	6.1
顺铂	969	9.2
伊立替康	263	2.5
紫杉醇	859	8.1
合计	10572	100

表2 支持度筛选

排名	ADR	例次	排名	ADR	例次
1	恶心	3046	9	皮疹	250
2	呕吐	2640	10	非特异性食欲异常	217
3	骨髓抑制	1515	11	粒细胞减少	214
4	白细胞减少	1066	12	胸闷	214
5	局部麻木	541	13	乏力	210
6	腹泻	525	14	肝功能异常	185
7	瘙痒	285	15	血小板减少	165
8	消化系统反应	280			

2.2.2 交叉列联表的建立

选取支持度 ≥ 0.01 不良反应和支持度 ≥ 0.01 的药品数据,并建立交叉列联频数表,见表3。

运用Clementine 14.1软件中的K-Means聚类模型对交叉列联频数表进行聚类,得出聚类结果。详见表3。

表3 ADR和药品的交叉列联频数表

ADR	奥沙利铂	表柔比星	多西他赛	氟尿嘧啶	卡培他滨	顺铂	伊立替康	紫杉醇
恶心	1377	63	132	787	121	405	44	117
呕吐	1169	66	105	679	122	342	54	103
骨髓抑制	471	45	235	269	82	167	61	185
白细胞减少	370	39	106	213	51	118	28	141
局部麻木	485	0	7	12	16	4	0	17
皮疹	157	2	16	17	21	6	9	22
胸闷	102	0	59	7	4	4	3	35
血小板减少	77	2	6	16	23	16	6	19
腹泻	163	5	40	140	64	21	65	27
瘙痒	220	1	18	10	10	4	3	19
消化系统反应	129	7	18	69	16	28	2	11
非特异性食欲异常	82	3	13	72	16	16	5	10
粒细胞减少	63	12	18	44	16	27	4	30
乏力	99	6	19	30	18	15	6	17
肝功能异常	70	5	17	46	16	17	1	13

2.2.3 按照药品聚类

按照药品进行聚类,共聚为5类,包括:第一类包括1种药品,第二类包括2种药品,第三类包括1种药品,第四类包括2种药品,第五类包括2种药品。最大聚类大小占25%,最小聚类大小占12.5%,二者比例为2:1。对预测变量的重要性进行排序,瘙痒的重要性最高,说明其在聚类过程中的运算比重较大。局部麻木、非特异性食欲异常、乏力的重要性依次减少。对聚类结果进行分析,结果分为五大类:1)奥沙利铂;2)表柔比星和伊立替康;3)氟尿嘧啶;4)紫杉醇和多西他赛;5)卡培他滨和顺铂。详见表4及表5。

2.2.4 按照ADR聚类

按照ADR进行聚类,共聚为2类。第一类包括2种ADR,第二类包括13种ADR。最大聚类大小占85.7%,最小聚类大小占14.3%,二者比例为6:1。对预测变量的重要性进行排序,顺铂的重要性最高,说明其在聚类过程中的运算比重较大。氟尿嘧啶、奥沙利铂重要性依次减少;伊立替康对预测变量的重要性最低。对聚类结果进行分析,ADR分为两大类:1)恶心和呕吐;2)骨髓抑制、白细胞减少、粒细胞减少、腹泻、血小板减少、肝功能异常和胸闷等。详见表4及表5。

表4 药品聚类结果

项目	奥沙利铂	表柔比星	多西他赛	氟尿嘧啶	卡培他滨	顺铂	伊立替康	紫杉醇
重要性	0.84	0.62	0.36	0.99	0.66	1.00	0.22	0.25
聚类结果	聚类-1	聚类-2	聚类-4	聚类-3	聚类-5	聚类-5	聚类-2	聚类-4

表5 药品不良反应聚类结果

ADR	重要性	聚类结果	ADR	重要性	聚类结果
恶心	0.44	聚类-1	消化系统反应	0.60	聚类-2
呕吐	0.46	聚类-1	非特异性食欲异常	0.91	聚类-2
骨髓抑制	0.39	聚类-2	粒细胞减少	0.31	聚类-2
皮疹	0.60	聚类-2	胸闷	0.41	聚类-2
局部麻木	0.98	聚类-2	白细胞减少	0.41	聚类-2
腹泻	0.22	聚类-2	肝功能异常	0.69	聚类-2
瘙痒	1.00	聚类-2	血小板减少	0.42	聚类-2
乏力	0.91	聚类-2			

2.3 按严重不良反应挖掘的K-Means聚类结果

2.3.1 支持度筛选

按频率将原患疾病及严重药品不良反应排序,选取支持度 ≥ 0.01 并筛选出发生率 ≥ 0.01 的不良反应,详见表6。表6中的数字表示某药品造成某不良反应的次数。

2.3.2 交叉列联表的建立

选取支持度 ≥ 0.01 严重药品不良反应和支持度 ≥ 0.01 的药品数据,建立交叉列联频数表。运用Clementine 14.1软件中的K-Means聚类模型对交叉列联频数表进行聚类,得出聚类结果,详见表7。

表6 支持度筛选

排名	ADR	例次	排名	ADR	例次
1	骨髓抑制	906	7	腹泻	75
2	白细胞减少	492	8	粒细胞缺乏症	72
3	呕吐	113	9	肝功能异常	70
4	血小板减少	98	10	胸闷	39
5	粒细胞减少	98	11	过敏性休克	37
6	恶心	92	12	呼吸困难	33

表7 严重ADR和药品的交叉列联频数表

严重ADR	奥沙利铂	表柔比星	多西他赛	氟尿嘧啶	卡培他滨	顺铂	伊立替康	紫杉醇
骨髓抑制	277	29	147	168	45	91	34	115
白细胞减少	151	8	58	94	21	55	13	92
呕吐	59	2	11	15	9	6	3	8
血小板减少	40	1	4	13	12	12	4	12
粒细胞减少	23	3	6	19	5	12	13	28
恶心	47	1	9	13	11	5	1	5
腹泻	21	2	10	13	12	3	10	4
粒细胞缺乏症	12	3	15	14	4	8	4	12
肝功能异常	29	2	5	14	7	4	0	9
呼吸困难	21	0	7	0	0	0	0	4
过敏性休克	24	0	4	2	0	1	0	6
胸闷	21	0	9	1	0	2	0	6

2.3.3 按照药品聚类

使用K-Means聚类模型,按照药品进行聚类共聚为3类。第一类包括1种药品,第二类包括4种药品,第三类包括3种药品。最大聚类大小占40%,最小聚类大小占10%,二者比例为4:1。对预测变量的重要性进行排序,过敏性休克的重要性最高,说明其在聚类过程中的运算比重较大。呕吐、恶心、呼吸困难重要性依次减少。对聚类结果进行分析,结果分为三大类:1)奥沙利铂;2)表柔比星、卡培他滨、顺铂和伊立替康;3)紫杉醇、氟

尿嘧啶和多西他赛,详见表8及表9。

2.3.4 按照ADR聚类

按照ADR进行聚类共聚为4类:第一类包括1种ADR,第二类包括6种ADR,第三类包括1种ADR,第四类包括2种ADR。最大聚类大小占66.7%,最小聚类大小占8.3%,二者比例为8:1。对预测变量的重要性进行排序,多西他赛的重要性最高,说明其在聚类过程中的运算比重较大。氟尿嘧啶,表柔比星难重要性依次减少;卡培他滨对预测变量的重要性最低。

对聚类结果进行分析,结果分为4大类:1)骨髓抑制;2)呕吐,血小板减少,恶心,粒细胞缺乏症,肝功能异常和胸闷;3)白细胞减少;4)粒细胞减少和腹泻。详见表8和表9。

表8 药品聚类结果

项目	奥沙利铂	表柔比星	多西他赛	氟尿嘧啶	卡培他滨	顺铂	伊立替康	紫杉醇
重要性	0.70	0.84	1.00	0.85	0.39	0.77	0.69	0.72
聚类结果	聚类-1	聚类-2	聚类-3	聚类-3	聚类-2	聚类-2	聚类-2	聚类-3

表9 药品不良反应聚类结果

ADR	重要性	聚类结果	ADR	重要性	聚类结果
恶心	0.70	聚类-2	血小板减少	0.50	聚类-2
呕吐	0.96	聚类-2	过敏性休克	1.00	-
骨髓抑制	0.65	聚类-1	粒细胞减少	0.15	聚类-4
胸闷	0.59	聚类-2	粒细胞缺乏症	0.54	聚类-2
呼吸困难	0.69	-	白细胞减少	0.52	聚类-3
腹泻	0.22	聚类-4	肝功能异常	0.55	聚类-2

3 讨论

3.1 聚类关系分析

聚类关系结果显示,卡培他滨和顺铂联合用药时,出现ADR的概率较高。表柔比星和伊立替康聚为一类,可以认为表柔比星和伊立替康联合用药发生ADR的可能性较大。多西他赛和紫杉醇聚为一类,可以认为多西他赛和紫杉醇联合用药发生ADR的可能性较大。

卡培他滨、顺铂、伊立替康和表柔比星联合用药时,出现严重ADR的概率较高。多西他赛、氟尿嘧啶和紫杉醇联合用药时,出现严重ADR的概率较高。可能是临床用药时氟尿嘧啶与多西他赛、紫杉醇联合用药的情况较为普遍,也可能是氟尿嘧啶与其他二者联合用药更容易出现药品间的相互作用。

在一般药品不良反应的聚类运算过程中,瘙痒、顺铂的重要性最高,说明其在聚类过程中的运算比重较大。在严重药品不良反应的聚类运算过程中,过敏性休克、多西他赛的重要性最高,说明其在聚类过程中的运算比重较大。

总结以往的文献,对本研究选用的8种药物药

品不良反应进行总结:奥沙利铂药品不良反应皮肤症状和呼吸系统症状发生率高。表柔比星容易外渗漏于皮下组织,轻则引起局部红肿热痛,重则可引起蜂窝组织炎和组织坏死或腐烂,甚至可造成器官性功能障碍。多西他赛可发生较严重的过敏反应,特征为低血压与支气管痉挛。氟尿嘧啶常见不良反应为骨髓抑制和胃肠道反应;偶见口腔黏膜炎、小脑共济失调、色素沉着、甲床变黑等,长期应用可致神经系统毒性。卡培他滨最常见毒副反应为消化道症状,主要为纳差,其次为恶心、呕吐;最严重的毒副反应为手足综合征,表现为疼痛、手掌和足底皮肤皲裂、斑样脱皮。顺铂不良反应主要有骨髓抑制、肾毒性、神经毒性、胃肠道毒性反应等。伊立替康主要毒副反应是胃肠道不良反应和骨髓抑制,其中包括不可预测的严重腹泻。紫杉醇注射液常见不良反应有胃肠道反应、神经毒性、骨髓抑制、过敏反应等。

3.2 临床合理用药

随着胃癌患者对于生存质量水平要求的不断提升,在肿瘤的治疗过程中,不能单一的考虑药品的治疗有效性。抗肿瘤药物引起ADR较多的原因

是该类药物选择性差,在抑制或杀伤肿瘤细胞的同时,对机体的某些正常功能造成损害^[5]。因此,医护人员应多关注抗肿瘤药物的不良反应,做到预防为主,及时处理^[6]。由于抗肿瘤药物作用机理的特殊性,要求医护人员在患者用药过程中时刻注意药品不良反应的发生情况,保障治疗的顺利进行。在多种化疗方案中,明确化疗药物的作用特点,才能最大限度地发挥抗肿瘤药物的治疗作用。

参考文献:

- [1] Martel C, Ferlay J, Franceschi S, et al. Global Burden of Cancers Attributable to Infections in 2008: A Review and Synthetic Analysis[J]. *The Lancet Oncology*, 2012, 13(6): 607-615.
- [2] 季加孚. 我国胃癌防治研究三十年回顾[J]. *中国肿瘤临床*, 2013, 40(22): 1346-1351.
- [3] 薛薇. 基于SPSS Modeler的数据挖掘[M]. 北京: 中国人民大学出版社, 2014.
- [4] 吴森. 基于数据挖掘技术的图书馆借阅记录分析研究[D]. 西安: 西安电子科技大学, 2011. 47-50.
- [5] 蔡和平, 王卓, 黄瑾, 等. 我院 108 例抗肿瘤药致不良反应报告分析[J]. *中国药房*, 2010, 21(34): 3235-3237.
- [6] 孙青, 张海林, 罗俊, 等. 937例药物不良反应报告分析[J]. *西北药学杂志*, 2016, 31(1): 96-98.

(收稿日期 2016年5月31日 编辑 范玉明)