

## 乳腺癌(乳岩)古今文献用药规律数据挖掘

刘德果<sup>1</sup>,李姿蓉<sup>2</sup>,刘凯<sup>3</sup>,陈杏元<sup>3</sup>,方颖<sup>3</sup>,范虹桥<sup>3</sup>,胡金辉<sup>3\*</sup>

(1.湖南中医药大学,湖南长沙410208;2.湖南省中医药研究院附属医院,湖南长沙410006;

3.湖南中医药大学第一附属医院,湖南长沙410007)

**[摘要]** 目的 为全面系统地了解古今中医文献中乳腺癌(乳岩)的用药规律,总结有效治疗措施指导临床实践。方法 将相关中医文献经过数据整理后建立乳腺癌文献信息标准数据库集,并在数据整理的基础上采用 Clementine 12.0 进行用药关联规则分析,得出古今中医文献中关于乳腺癌的用药规律,并运用 SPSS 19.0 进行文献中所用药物的使用频数/频率进行聚类分析。结果 治疗乳腺癌出现频次排名前五的药物类别由多到少依次为理气药>活血药>补阳药>滋阴药>化痰药。理气药常与活血药、补阳药、滋阴药、化痰药合用。结论 乳腺癌治疗以理气化瘀药为主,联合应用化痰药、补阳药、滋阴药的用药规律切合临床实际,值得推广应用。

**[关键词]** 乳腺癌;乳岩;补肾活血汤;关联规则;聚类分析;数据挖掘

**[中图分类号]**R273 **[文献标识码]**A **[文章编号]**doi:10.3969/j.issn.1674-070X.2016.04.019

## Data Mining for Medication Rules of Breast Cancer (Mammary Cancer) in Ancient and Modern Literatures

LIU Deguo<sup>1</sup>, LI Zirong<sup>2</sup>, LIU Kai<sup>3</sup>, CHEN Xingyuan<sup>3</sup>, FANG Ying<sup>3</sup>, FAN Hongqiao<sup>3</sup>, HU Jinhui<sup>3\*</sup>

(1.Hunan University of Chinese Medicine, Changsha, Hunan 410208, China; 2.Hunan Academy of Chinese Medicine, Changsha, Hunan 410006, China; 3. The First Affiliated Hospital of Hunan University of Chinese Medicine, Changsha, Hunan 410007, China)

**[Abstract]** **Objective** To comprehensively understand the medication rules of breast cancer (mammary cancer) in ancient and modern literatures. The clinical practice can be guided by summarized effective treatment measures. **Methods** The standard database of breast cancer information was built by collecting the related Chinese medicine literatures with data integration. On the basis of the datacollecting, the association rules were analyzed by using Clementine 12.0, and the medication rules of breast cancer were obtained. The using frequency of medicine were made clustering analysis by SPSS 19.0. **Results** The former five frequency of medicine was Qi-regulating drug>circulating drug> Yang-toning drug>norishing Yin drug>expectorant druguse for literature of the drugs used in treatment of breast cancer. **Conclusion** Qi-regulating and removing blood stasis class drug treat breast cancer for the medicine primarily, combining with expectorant drug, Yang-tonifying drug and norishing Yin drug. The medication rules conform to the clinical practice, and are worthy of popularization and application.

**[Keywords]** breast cancer; mammary cancer; norishing kidney and circulating blood decoction; literature research; association rules; clustering analysis; data mining

据国际抗癌联盟的最新统计显示:乳腺癌是世界范围内女性发病率最高的恶性肿瘤,占女性新发恶性肿瘤的31%,其五年生存率仅为46%,死亡率仅次于肺癌居第2位<sup>[1]</sup>。而根据监测、流行病学和最终结果数据库的数据显示,在80年代美国的乳腺癌

发病率是106.07/10万,到了2012年,上升到了146.02/10万,至少增长了35%<sup>[2]</sup>。2012年乳腺癌全球范围内的医疗花费达1370亿美元,占全球医疗支出的3.2%<sup>[3]</sup>。目前我国乳腺癌的防治及其术后转移治疗的形势十分严峻,我国的乳腺癌发病率为

**[收稿日期]**2015-10-10

**[作者简介]**刘德果,男,在读硕士研究生,方向:中医外科学。

**[通讯作者]**\*胡金辉,女,主任医师,E-mail:502522782@qq.com。

63.6 /10 万,五年生存率仅为 34.7%,现阶段乳腺癌已经成为世界各国必须共同面对的巨大生命威胁<sup>[4]</sup>。

与乳腺癌防治的严峻形势相对应的是,我国古今中医文献中记载了较为完整的乳腺癌诊疗理论体系以及较为丰富的临床经验,对有效防治乳腺癌、提高乳腺癌患者的生活质量具有重要的作用。然而不同医家,不同中医文献在乳腺癌诊疗方法的定量化、规范化和客观化等方面存在着许多差异,形成了中医文献“资源丰富,数据贫乏”的现状。研究组以古今中医丰富的乳腺癌相关文献资源为依托,通过对文献资料的数据化处理,建立中医文献信息数据库,引入数据挖掘技术,对广泛的中医数据运用 SPSS 19.0、Clementine 12.0 等统计软件分析,挖掘乳腺癌的中医用药规律,现将方法与结果报道如下。

### 1 资料与方法

#### 1.1 资料

文献数据来源于《伤寒杂病论》《肘后备急方》《诸病源候论》《外台秘要》《备急千金要方》《外科精要》《外科正宗》《中西医结合防治肿瘤》等 61 部古今中医文献电子书。

#### 1.2 方法

1.2.1 检索策略 检索词为“乳腺癌”or“乳癌”or“乳岩”。

1.2.2 建立乳腺癌文献信息数据库 筛选出的中医乳腺癌相关文献,将文献当中的医案及用药等数据经专人仔细研读,依循编号、文献名称、作者、成书年代、病因、病机、症状、方剂、用药、剂量、剂型、服用方法、功效、禁忌等结构采集文献信息,建立信息丰富的乳腺癌原始数据库。

文献数据的整理总体上分为 3 个步骤:(1) 数据选取和分类;(2)数据预处理;(3)数据转换。如图 1 所示。

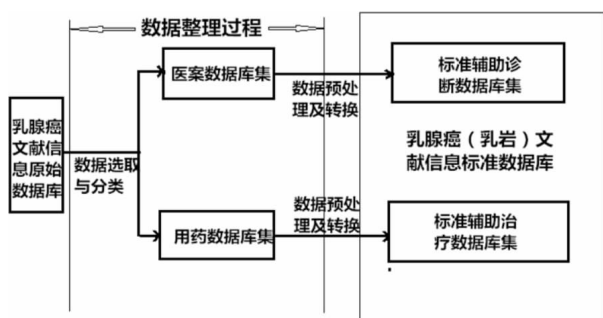


图 1 数据库整理流程图

#### (1)数据选取与分类

将乳腺癌文献中的原始数据形式经过适当的处理,在保持文献数据原貌的前提下将其数字化,以便于其适合计算机分析,同时方便文献数据的入库,也为其后的数据挖掘奠定一定的基础。综上所述,研究组将中医文献原始数据库分为诊断数据库集和用药数据库集,图 2-3 针对乳腺癌诊断和用药数据库集做出了整体架构。

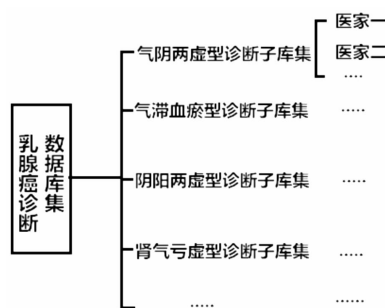


图 2 乳腺癌诊断数据库集的架构

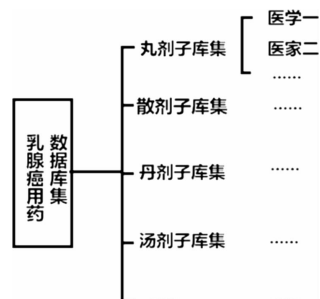


图 3 乳腺癌用药数据库集的架构

#### (2)文献数据预处理

参照《中医方剂大词典》<sup>[5]</sup>《中华本草》<sup>[6]</sup>《中药大辞典》<sup>[7]</sup>《中华人民共和国国家标准·中医临床诊疗术语》<sup>[8]</sup>《中医诊断学》<sup>[9]</sup>《中国科学技术史·度量衡卷》<sup>[10]</sup>对方剂、药名、症状、药物剂量进行规范化处理。

#### (3)文献数据转换

对规范化以后的临床症状、用药名称、用药剂量及服用方法等数据按一定顺序进行编码转变为数字化的表现形式,表 1 给出了乳腺癌部分辨证名称的编码序列,并且把图 2-3 中的诊断/用药数据库集转化为条件属性、决策属性的形式。若为诊断数据库集,条件属性可为临床症状或体征,而决策属性可为辨证后的症状名;若为用药数据库集,则条件属性可为中草药名称或者用药的剂量,决策属性可为方名或者功效,经过这一步骤后便构建了乳腺癌(乳岩)

的文献信息标准数据库集。

表1 乳腺癌部分辨证名称编码序列结构图

辨证分型	对应代码
气滞血瘀	1
肾气亏虚	2
阴阳两虚	3
...	...

### 1.3 统计学分析

数据的数字化利用二分类法,将规范化后的症状、中药按照出现标为“1”、不出现“0”录入 Excel 表格,进行数据的整理、频数统计、采用 SPSS 19.0 软

件进行药物的聚类分析、Clementine 12.0 进行药物关联规则分析,在此基础上结合专业知识,对统计结果进行分析评价,挖掘乳腺癌(乳岩)古今文献中的用药规律。

## 2 结果

### 2.1 不同时期乳腺癌(乳岩)高频使用药物类别结果

乳腺癌治疗中使用药物类别排名前5的类别由多到少依次为理气药>活血药>补阳药>滋阴药>化痰药。高频使用药物类别频数与频率在不同时期变化较大。见表2。

表2 不同时期乳腺癌(乳岩)高频使用药物类别结果

(次,%)

类别	唐宋以前		金元时期		明清时期		近现代	
	频数	频率	频数	频率	频数	频率	频数	频率
理气药	471	35.12	52	3.87	231	17.23	587	43.78
活血药	362	28.64	84	6.65	386	30.53	432	34.18
补阳药	213	21.98	147	15.17	284	29.31	325	33.54
滋阴药	198	17.63	134	11.93	517	46.04	274	24.40
化痰药	86	13.56	114	17.98	124	19.56	310	48.90

### 2.2 乳腺癌(乳岩)用药聚类分析结果

聚类分析为二分类时,理气药与其他药物类型明显区分;当为三分类时,理气药与活血药可聚为一类;而补阳药与滋阴药为一类;化痰药聚为一类。结果见表3、图4。

表3 乳腺癌(乳岩)用药聚类分析结果

序号(多→少)	组合
1	理气药组+活血药组
2	理气药组+补阳药组
3	活血药组+补阳药组
4	理气药组+滋阴药组
5	理气药组+化痰药组
6	活血药组+滋阴药组
7	活血药组+化痰药组
8	滋阴药组+补阳药组
9	补阳药组+化痰药组
10	滋阴药组+化痰药组

### 2.3 乳腺癌(乳岩)用药关联规则分析结果

从文献中筛选置信度大于0.6的中药类型组合,分别是“理气药+活血药”,“理气药+补阳药”,“活血药+补阳药”,“理气药+滋阴药”,“理气药+化痰药”,“活血药+滋阴药”,“活血药+化痰药”,“滋阴药+补阳药”,“补阳药+化痰药”,“滋阴药+化痰

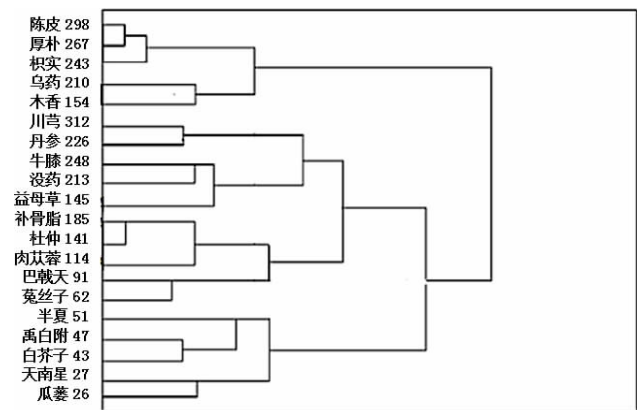


图4 中药类型组合聚类分析图

药”。且理气药组与活血药组关联度最高,理气药组与补阳药组次之,化痰药可运用于乳腺癌(乳岩)的各个药组之中。见表4。

## 3 讨论

(1)不同年/朝代理气药和活血药是医家们使用频率最多的两类中药(表3-4),由此可得出不同年/朝代的医家们把气滞血瘀视作乳腺癌产生的最主要原因,亦符合乳腺癌(乳岩)的病因病机以气滞血瘀、癌毒瘀于体内为要,此二者通常是相互兼夹。又因血为气之母,气为血之帅,气血之间的相互关系决定了

表4 乳腺癌(乳岩)用药关联规则分析结果

序号	A类药	B类药	支持度 A→B	置信度 A→B
1	理气药组	活血药组	0.865	0.976
2	理气药组	补阳药组	0.719	0.976
3	活血药组	补阳药组	0.746	0.934
4	理气药组	滋阴药组	0.714	0.933
5	理气药组	化痰药组	0.637	0.875
6	活血药组	滋阴药组	0.637	0.875
7	活血药组	化痰药组	0.637	0.750
8	滋阴药组	补阳药组	0.785	0.725
9	补阳药组	化痰药组	0.635	0.695
10	滋阴药组	化痰药组	0.695	0.695

此两种病理因素的相互转化关系,气滞、血瘀是乳岩病发病的病理中心与关键环节,常与其他病理因素兼夹为病,如气血痰相兼,治疗以活血化瘀为主,兼以祛痰。由此可见,治疗乳岩病若能紧紧抓住气滞血瘀,则此病可解其半。

(2)唐宋以前的医家们认为乳腺癌的主要病因病机为气滞血瘀、积毒伤于乳络,而素体阳虚或阴虚则为次要因素;金元时期的医家们则认为肾精亏虚,脏腑虚弱,阴阳失调,导致癌毒侵袭为乳腺癌的主要病因病机,认为乳腺癌是本虚标实,正气虚为其本,瘀血邪毒互结为其标,以扶正祛邪、标本兼治为治则,其次才为气滞和血瘀,同时痰湿阻滞气机,经络不通等亦为乳腺癌的病因病机;明清时期的医家们更偏重于机体肾阴亏损的病因病机的变化,并且血瘀、气滞、痰湿等病因致癌的理论也日益成熟;现代治疗乳腺癌以理气化瘀,补益虚损为主要治法,辅以燥湿化痰等治法。这表明乳腺癌(乳岩)的诊疗体系及治疗理论在历史的传承中得到了进一步的发展。

(3)人体五脏六腑相互关联,一荣俱荣,一损俱损。在中医文献中乳腺癌在脏腑辨证中多责之于肝、肾、脾胃等脏腑,又与其他脏腑息息相关。肾的先天精气、脾胃的后天水谷之气、肝的藏血与输调气机均对乳腺癌的发生、发展有着密切的关系,医家们大多比较灵活地运用脏腑辨证将相关脏腑一一细心诊察,调治,有则治之,无则护之。而人体的气血津液亦参与了病理因素的形成,人未病时是人身之正气,人已病时既为抗邪之本又为病邪之始,乳腺癌的气血津液辨证亦有重要地位。研究组针对文献中乳腺癌患者用药以理气药、活血药居多,从古代中医文献中

挖掘出补肾活血汤并运用于临床多年,疗效明显。

(4)在文献中研究组发现,化痰药,补阳药,滋阴药等亦占有重要部分。在乳腺癌后期,正气已虚,余毒未清,体内残留的癌灶、毒素在机体正气虚弱的情况下随时可能复发,机体病理代谢产物湿、瘀、毒等在体内亦有不同程度的蓄积。故此时应采用“扶正为主,祛邪为辅,病症结合,标本兼治”的治则,调整病者机体阴阳、气血、脏腑功能,恢复正气,清解余毒。这对现代临床中治疗乳腺癌亦有重要启示。

(5)研究中运用了文献构造<sup>[11-12]</sup>的“属性约简加改进值约简”算法在乳腺癌的疾病诊断和用药配伍等数据库集进行更深层次的数据挖掘,以期挖掘出更符合现今社会临床需求的古今中医文献中关于乳腺癌(乳岩)的诊疗手段(规则)。目前研究组用药多从古代文献化裁古方并结合现代药理学研究加入确有疗效的药物,不拘泥于中医文献又不菲薄今人研究,去其糟粕,取之精华,为己所用。对古今中医文献进行数字化研究,是中医在继承与创新之路上进行的有益探索。

#### 参考文献:

- [1] Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013[J].CA Cancer J Clin, 2013,6(1):11-30.
- [2] DeSantis C, Ma J, Bryan L, et al. Breast cancer statistics,2013[J].CA Cancer J Clin,2014,64(1):52-62.
- [3] 何俊安.乳腺癌的患病因素与中医未病先防的相关性研究[J].现代中西医结合杂志,2014,23(14):1 551-1 553.
- [4] 薛娜,杨国旺,王笑民.乳腺癌中医证候研究现状及思考[J].中医杂志,2014,55(2):168-171.
- [5] 彭怀仁.中医方剂大词典[M].北京:人民卫生出版社,2005.
- [6] 国家中医药管理局《中华本草》编委会.中华本草[M].上海:上海科学技术出版社,1999.
- [7] 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.
- [8] 国家技术监督局.GBT 16751.2-1997 中医临床诊疗术语 [S].北京:中国标准出版社,2004:1-88.
- [9] 朱文峰.中医诊断学[M].北京:中国中医药出版社,2007.
- [10] 丘光明.中国科学技术史·度量衡卷 [M].北京:科学出版社,2001:578.
- [11] 蒋新军,陈燕.中医治未病学说文献探源[J].湖南中医药大学学报,2013,33(11),10-12.
- [12] 金力,王宗殿,阙红星,等.数据挖掘在中医诊疗规则提取中的应用研究[J].时珍国医国药,2013,24(4):1 015.

(本文编辑 贺慧斌)