

胡珊珊, 厉玉鹏. 儿童过敏性紫癜用药规律的数据挖掘[J]. 中华医学图书情报杂志, 2018, 27(6): 54-58.

DOI: 10. 3969/j. issn. 1671-3982. 2018. 06. 009

· 信息组织与信息服务 ·

儿童过敏性紫癜用药规律的数据挖掘

胡珊珊¹, 厉玉鹏²

[摘要] 介绍数据挖掘技术的常用算法, 利用数据挖掘的关联规则算法挖掘儿童过敏性紫癜的用药模式, 发现治疗过程中的核心用药以及药品间的关联规则, 总结过敏性紫癜儿童用药的特点和潜在规律, 挖掘结果显示双嘧达莫片与盐酸西替利嗪糖浆、盐酸西替利嗪糖浆与西咪替丁注射液在儿童过敏性紫癜治疗中具有较高相关性, 为医疗大数据背景下过敏性紫癜儿童用药决策及合理用药探讨提供循证医学支持。经专家证实, 挖掘结果与实际的临床用药情况相符, 具有实用价值和指导意义。

[关键词] 医疗大数据; 数据挖掘; 儿童过敏性紫癜; 关联规则; 决策支持

[中图分类号] R725. 5; R-058

[文献标志码] A

[文章编号] 1671-3982(2018)06-0054-05

Data mining of drug use rules in treatment of allergic purpura children

HU Shan-shan¹, LI Yu-peng²

(1. Jining Medical College Medical Information Engineering School, Rizhao 276826, Shandong Province, China; 2. People's Hospital of Rizhao, Rizhao 276826, Shandong Province, China)

[Abstract] After the background of medical big data, and the concepts and commonly used algorithms of data mining technology were described, the drug use methods in treatment of allergic purpura children were mined according to the association rules algorithm of data mining, which showed that the methods for use of key drugs, association rules between different drugs, characteristics and potential rules of drugs used in treatment of allergic purpura children can provide evidence for policy decision and rational use of drugs in treatment of allergic purpura children. The mining results verified by experts are in line with the practical use of drugs and are thus of practical values.

[Key words] Medical Big data; Data mining; Allergic purpura children; Association rules; Decision-making support

随着医疗信息化程度日益加深, 医疗数据大量积累, 医疗大数据时代已经到来^[1-2]。医疗大数据是对患者在医疗活动中产生的全体数据的集合, 包括患者个人信息、诊疗信息、电子病历、用药数据、医

学影像数据等^[3]。医疗大数据为医疗卫生事业的发展带来了巨大的机遇, 也为循证医学的发展提供了宝贵的数据资源。为促进医疗大数据的发展, 我国陆续出台了多项相关政策, 将医疗大数据发展纳入国家大数据发展战略, 全面推进医疗大数据的挖掘和应用^[4-5]。如何科学有效地对医疗大数据进行挖掘与分析, 为临床实践提供全面、可靠的数据支持, 推动强调证据为基础的循证医学的发展, 促进医疗大数据在医学领域的应用与研究, 成为当前医疗行业发展的必然趋势^[6-7]。

儿童的各个器官都处在生长发育的特殊阶段, 肝肾解毒功能较差, 容易出现用药安全问题^[8]。据

[基金项目] 济宁医学院校级科研计划项目“大数据时代针对儿童安全用药的数据挖掘技术研究”(JY2015KJ020); 济宁医学院教师科研扶持基金“面向基层医院的健康医疗大数据挖掘和应用研究”(JY2017KJ056)

[作者单位] 1. 济宁医学院医学信息工程学院, 山东日照 276826; 2. 日照市人民医院, 山东日照 276826

[作者简介] 胡珊珊(1980-), 女, 山东日照人, 硕士, 讲师, 参编教材 3 部, 发表论文 9 篇, 研究方向为医学信息技术、医疗大数据挖掘。

调查,我国每年约 30 000 名儿童由于用药不当导致耳聋,约 7 000 名儿童因此失去生命^[9]。因此,充分挖掘儿童用药大数据,发现其中潜在的规律,可为儿科医生临床用药提供决策支持。本文尝试将数据挖掘技术应用于儿童过敏性紫癜用药模式分析,为儿童用药决策以及合理用药提供循证医学支持。

1 数据挖掘技术

1.1 概念和挖掘流程

数据挖掘是指通过对大量有噪声、不完整的模糊数据进行探索和分析,发现数据有潜在价值的模式和知识^[10-11]。数据挖掘的流程包含数据收集、数据预处理、数据挖掘、结果解释与评价^[12]。

1.2 挖掘算法

数据挖掘有多种算法,本文使用关联规则算法进行挖掘,关联规则算法包含频繁项集和关联规则 2 个核心概念。频繁项集是指经常一起使用的药品或药品组合,包含 k 种药品的项集称为 k -项集,频繁项集要满足最小支持度,即药品出现的频率要高于某一阈值;关联规则是蕴含关系 $A \rightarrow B$,表示如果使用了药品 A ,就会与使用药品 B 产生关联。强关联规则要满足最小支持度和最小置信度,有效的关联规则要满足提升度 > 1 。支持度(Support)、置信度(Confidence)和提升度(Lift)的公式如下^[13-14]:

$$\text{Support}(A \rightarrow B) = P(A \cup B)$$

$$\text{Confidence}(A \rightarrow B) = P(B/A)$$

$$\text{Lift}(A \rightarrow B) = P(B/A) / P(B)$$

2 儿童过敏性紫癜用药数据挖掘

本文利用数据挖掘技术对儿童用药模式进行挖掘。过敏性紫癜是儿童多发的一种血管炎疾病,主要表现为皮肤紫癜,伴有关节肿痛、腹痛、肾脏损害和消化道出血等症状^[15-16]。通过对儿童过敏性紫癜药品使用频率、药品之间的关联性进行分析,发现儿童用药的特点及规律。

2.1 数据收集

收集 2012 年 1 月~2018 年 5 月日照市某市级公立医院收治的年龄介于 2~14 岁的 505 例(男孩 309 例,女孩 196 例)过敏性紫癜患儿用药数据。

2.2 数据预处理

医疗大数据经常存在噪声、不完整和不一致的问题,因此在数据挖掘前要对儿童过敏性紫癜用药

数据进行预处理。

2.2.1 隐私保护和属性提取

医疗大数据中包含多项患者的隐私信息,必须进行保护。首先对数据库中的数据进行匿名化处理,隐藏与患者隐私相关的属性,如姓名、家族病史等信息;然后根据数据挖掘的目标,在数据库中进行属性选择,提取患者的病历号、医嘱内容、医嘱时间等属性,忽略其他无关的属性。

2.2.2 数据清洗

对数据库中的医嘱数据进行清洗。除去临床上常用的药物溶媒以及护理、禁食等医嘱内容,对不同时间段多次使用某种药品造成的数据冗余进行消除重复处理,填补医嘱数据中的缺失,删除医嘱数据中的异常值,提高数据的完整性。

2.2.3 药品名规范和编码

由于医院数据库中的药品名称不规范,根据《中华人民共和国药典》将药品名规范统一化。如将“西替利嗪糖浆”统一为“盐酸西替利嗪糖浆”,“孟鲁司特咀嚼片”统一为“孟鲁司特钠咀嚼片”,保证药品名的一致性。最后,用 d_1, d_2, \dots, d_n 对药品名进行编码,其中,“ n ”表示药品种数,能够在某种程度上简化挖掘过程。

2.3 数据挖掘

在 R 语言中实现关联规则 Apriori 算法,挖掘儿童过敏性紫癜用药模式。为便于 R 语言接收,需要把数据另存为 csv 格式,通过 read.csv() 函数读入 R 中,用 data 表示读入的待挖掘数据。

2.3.1 用药统计汇总

通过 summary() 函数对 data 统计汇总,得到最小值、最大值、四分位数和均值等统计信息。语句为 summary(data)。

2.3.2 发现频繁项集

通过 eclat() 函数,设置支持度阈值为 20%,最大项数为 1,找出用药数据的频繁 1 项集。语句为 eclat(data, parameter = list(support = 0.2, maxlen = 1))。

通过 eclat() 函数,设置支持度阈值为 30%,最小项数为 2,最大项数为 5,找出频繁项集。语句为 eclat(data, parameter = list(support = 0.3, maxlen = 5, minlen = 2))。

2.3.3 发现关联规则

通过 apriori () 函数, 设置支持度阈值 70%, 置信度阈值 80%, 得到用药数据的关联规则。语句为 apriori (data, parameter = list (support = 0.7, confidence = 0.8))。

3 结果解释与评价

3.1 用药统计信息

儿童过敏性紫癜用药数据的统计信息见表 1。

表 1 儿童过敏性紫癜用药统计信息

统计量	统计值
最小值	2.000
第一四分位数	6.000
中位数	8.000
均值	7.604
第三四分位数	9.000
最大值	13.000

从表 1 可以看出, 在不考虑溶媒的条件下, 过敏性紫癜患儿在治疗过程中最少使用 2 种药品, 最多使用 13 种药品; 四分之一的患儿用药种数 ≤ 6 种, 一半的患儿用药种数 ≤ 8 种, 四分之三的患儿用药种数 ≤ 9 种, 而患儿平均用药种数为 7.604 种。

3.2 频繁项集

3.2.1 频繁 1 项集

利用函数 inspect () 察看支持度最高的前 10 项频繁 1 项集, 得到过敏性紫癜儿童最常用的 10 种药品 (表 2)。

表 2 过敏性紫癜儿童常用药品

药品	支持度/%
维生素 C 注射液	97.03
西咪替丁注射液	95.05
盐酸西替利嗪糖浆	89.11
双嘧达莫片	85.15
低分子肝素钠注射液	82.18
葡萄糖酸钙注射液	47.52
喜炎平注射液	37.62
注射用甲泼尼龙琥珀酸钠	35.64
注射用乳糖酸阿奇霉素	31.68
醋酸泼尼松片	27.72

从表 2 可知, 维生素 C 注射液、西咪替丁注射液、盐酸西替利嗪糖浆、双嘧达莫片和低分子肝素钠注射液应用最多。维生素 C 注射液主要用于紫癜的治疗, 西咪替丁注射液可以抗组胺、减轻腹痛, 盐酸西替利嗪糖浆主要是抗组胺, 双嘧达莫片用于阻止血小板聚集和血栓形成, 低分子肝素钠注射液可以抗凝血、预防血栓, 说明过敏性紫癜的治疗主要为对症支持, 抗组胺、抗凝治疗。

3.2.2 频繁项集

利用函数 inspect () 察看支持度最高的前 10 项频繁项集, 得到最常用的 10 种药品组合 (表 3)。

表 3 过敏性紫癜儿童常用药品组合

药品组合	支持度/%
维生素 C 注射液, 西咪替丁注射液	92.08
西咪替丁注射液, 盐酸西替利嗪糖浆	88.12
维生素 C 注射液, 盐酸西替利嗪糖浆	86.14
维生素 C 注射液, 西咪替丁注射液, 盐酸西替利嗪糖浆	85.15
西咪替丁注射液, 双嘧达莫片	83.17
维生素 C 注射液, 双嘧达莫片	82.18
盐酸西替利嗪糖浆, 双嘧达莫片	81.19
维生素 C 注射液, 低分子肝素钠注射液	80.20
西咪替丁注射液, 盐酸西替利嗪糖浆, 双嘧达莫片	80.20
维生素 C 注射液, 西咪替丁注射液, 双嘧达莫片	80.20

在这些药品组合中, 维生素 C 注射液联合西咪替丁注射液是减少紫癜和抗组胺, 西咪替丁注射液联合盐酸西替利嗪糖浆是抗组胺和抗过敏, 维生素 C 注射液联合盐酸西替利嗪糖浆是减少紫癜和抗组胺、抗过敏, 西咪替丁注射液联合双嘧达莫片是抗组胺、减轻腹痛和抗凝, 说明过敏性紫癜儿童可能与食物、微生物和药物等因素相关, 但没有确凿证据, 因此发病原因难以确定。过敏性紫癜患儿的治疗主要为对症处理, 补充维生素, 应用抗组胺药物, 腹痛时应用西咪替丁、抗凝药物、激素和免疫抑制剂, 抗生素应用较少。

3.3 关联规则

利用 inspect () 函数按提升度 > 1.04 筛选并察看关联规则, 得到 12 条药品关联规则 (表 4)。

表 4 过敏性紫癜儿童药品关联规则

药品关联规则(前项->后项)	支持度/%	置信度/%	提升度
双嘧达莫片 -> 盐酸西替利嗪糖浆	81.19	95.35	1.0700
盐酸西替利嗪糖浆 -> 双嘧达莫片	81.19	91.11	1.0700
盐酸西替利嗪糖浆 -> 西咪替丁注射液	88.12	98.89	1.0404
西咪替丁注射液 -> 盐酸西替利嗪糖浆	88.12	92.71	1.0404
盐酸西替利嗪糖浆,低分子肝素钠注射液 -> 西咪替丁注射液	73.27	100	1.0521
维生素 C 注射液,双嘧达莫片 -> 盐酸西替利嗪糖浆	78.22	95.18	1.0681
维生素 C 注射液,盐酸西替利嗪糖浆 -> 双嘧达莫片	78.22	90.80	1.0664
西咪替丁注射液,双嘧达莫片 -> 盐酸西替利嗪糖浆	80.20	96.43	1.0821
西咪替丁注射液,盐酸西替利嗪糖浆 -> 双嘧达莫片	80.20	91.01	1.0689
维生素 C 注射液,盐酸西替利嗪糖浆,低分子肝素钠注射液 -> 西咪替丁注射液	71.29	100	1.0521
维生素 C 注射液,西咪替丁注射液,双嘧达莫片 -> 盐酸西替利嗪糖浆	77.23	96.30	1.0807
维生素 C 注射液,西咪替丁注射液,盐酸西替利嗪糖浆 -> 双嘧达莫片	77.23	90.70	1.0652

对部分关联规则说明如下:

双嘧达莫片和盐酸西替利嗪糖浆联合使用的概率是 81.19%。如果患儿使用了双嘧达莫片,则有 95.35% 的可能使用盐酸西替利嗪糖浆;如果使用了盐酸西替利嗪糖浆,则有 91.11% 的可能使用双嘧达莫片。

盐酸西替利嗪糖浆和西咪替丁注射液联合使用的概率是 88.12%。如果患儿使用了盐酸西替利嗪糖浆,则有 98.89% 的可能使用西咪替丁注射液;如果使用了西咪替丁注射液,则有 92.71% 的可能使用到盐酸西替利嗪糖浆。

盐酸西替利嗪糖浆联合低分子肝素钠注射液和西咪替丁注射液的概率是 73.27%。如果患儿使用了盐酸西替利嗪糖浆和低分子肝素钠注射液,则有 100% 的可能使用西咪替丁注射液。

维生素 C 注射液联合双嘧达莫片和盐酸西替利嗪糖浆的概率是 78.22%。如果患儿使用了维生素 C 注射液和双嘧达莫片,则有 95.18% 的可能使用盐酸西替利嗪糖浆;如果使用了维生素 C 注射液和盐酸西替利嗪糖浆,则有 90.80% 的可能使用双嘧达莫片。

总的来看,双嘧达莫片与盐酸西替利嗪糖浆,盐酸西替利嗪糖浆与西咪替丁注射液在儿童过敏性紫癜治疗中具有较高的相关性。双嘧达莫片主要有抗凝的作用,盐酸西替利嗪主要用于抗组胺、抗过敏,西咪替丁主要是抗组胺、解痉、减轻腹痛,说明抗凝、

抗组胺、抗过敏是过敏性紫癜(皮肤型)的主要治疗方案,而抗组胺、抗过敏、护胃解痉是过敏性紫癜(腹型)的主要治疗方案。

此外,盐酸西替利嗪糖浆与低分子肝素钠注射液、西咪替丁注射液的关联广泛,3 种药品的合并用药可以抗组胺、抗凝、解痉,治疗过敏性紫癜腹痛。维生素 C 注射液与盐酸西替利嗪糖浆、双嘧达莫片的关联关系说明在减轻紫癜的基础上,抗组胺、抗凝是治疗儿童过敏性紫癜的主要思路。

经过与患儿所在医院的儿科主任医师沟通,以上挖掘结果与儿童过敏性紫癜的临床用药情况相符合,体现出了儿科医师治疗该病时的用药习惯和诊疗思路,在辅助临床给药、传承用药经验和促进用药规范等方面具有实用价值。

4 结语

随着医疗行业信息化的不断深入,医疗大数据的价值正日益显现。如何充分利用医疗大数据,挖掘医疗大数据背后隐藏的价值,提高医疗服务水平,成为医疗大数据时代的重要研究领域。

本文利用医疗大数据和数据挖掘技术,挖掘儿童过敏性紫癜的核心用药及药品间的关联规则,分析儿童用药的特点和规律,为指导儿童临床用药提供了循证医学支持。

但是需注意的是核心用药及用药规律仍要在临床实践中不断验证和评价,才能更好地发挥其在指导儿童用药方面的作用。

【参考文献】

[1] 郑西川,孙宇,陈霆,等.基于医疗大数据分析的临床电子病历智能化研究[J].中国数字医学,2016,11(11):61-64.

[2] 薛付忠.健康医疗大数据驱动的健康管理学理论方法体系[J].山东大学学报:医学版,2017,55(6):1-29.

[3] 陈素琼,王惠来,向天雨.医疗大数据应用现状研究[J].中国数字医学,2017,12(9):30-31.

[4] 戴明锋,孟群.医疗健康大数据挖掘和分析面临的机遇与挑战[J].中国卫生信息管理杂志,2017,14(2):126-130.

[5] 徐志祥,王莹.我国医疗行业大数据应用现状及政策建议[J].中国卫生信息管理杂志,2017,14(6):822-825.

[6] 陆易,黄正行,俞思伟,等.临床医疗大数据研究现状与展望[J].医疗卫生装备,2017,38(3):112-115.

[7] 牟忠林,王雅洁,陈娟,等.健康大数据在医疗卫生领域中的应用及挑战[J].海南医学,2017,28(2):173-176.

[8] 孙晓.儿童用药安全问题及对策研究[J].中国卫生产业,2017,14(4):23-24.

[9] 国家食品药品监督管理总局南方医药经济研究所.2016年儿

童用药安全调查报告白皮书[R].2016.

[10] 廖冬雪.数据挖掘技术在医学领域中的应用[J].财经界,2017(16):110-111.

[11] 郑娟,许建强.数据挖掘技术及其在医疗质量管理中的应用[J].医学信息学杂志,2018,39(3):70-73.

[12] 张雷,王云光.健康大数据挖掘方法研究综述[J].软件导刊,2018,17(3):1-3.

[13] 李昌兵,凌永亮,汪尔晶.基于兴趣度的Web访问用户关联规则挖掘[J].计算机工程与设计,2017,38(4):852-856.

[14] 马彦勤,武彤,邓焯堃.基于兴趣度度量的正负关联规则挖掘方法研究[J].计算机技术与发展,2018,28(5):38-41.

[15] 舒静,陈芳,何强,等.儿童过敏性紫癜合并胰腺炎临床特征及治疗随访分析[J].中国实用儿科杂志,2018,33(1):60-64.

[16] 马敏.小儿过敏性紫癜的临床诊治研究[J].临床研究,2018,26(2):94-95.

[收稿日期:2018-05-26]

[本文编辑:刘娜]