

主题模型在2型糖尿病治疗方案研究中的应用^Δ

杨帆^{1*}, 李海英¹, 许少杰¹, 邹兆重¹, 李忆琴², 陈画虹¹ (1. 中山市博爱医院药学部, 广东中山 528403; 2. 中山市博爱医院内分泌科, 广东中山 528403)

中图分类号 R969.3 文献标志码 A 文章编号 1001-0408(2017)23-3208-05
DOI 10.6039/j.issn.1001-0408.2017.23.11

摘要 目的: 探讨主题模型在2型糖尿病治疗方案研究中的应用。方法: 回顾性分析我院2012年1月—2016年6月650例2型糖尿病患者住院患者的临床数据, 采用MATLAB R2014a软件对患者的临床诊断、实验室指标和临床用药等数据进行转换、汇总及矩阵合并; 应用隐狄利克雷分配模型和作者主题模型提取主题概率值>0.1的典型主题, 并通过累计概率值>0.5的并发症对主题进行描述。结果: 共获得62个并发症词汇、16个异常实验室指标组和20个治疗方案。共挖掘出4个典型主题(前几项并发症的累计概率值分别为0.878 6、0.824 7、0.821 5、0.753 6, 主题概率值分别为0.336 4、0.277 3、0.203 5、0.117 6), 分别以周围神经病变、外周血管病变、脂代谢异常和微细血管病变为主要特征, 且异常实验室指标组符合上述特征。分布概率较高的并发症包括糖尿病周围神经病变(0.578 7)、高血压(0.363 1)、动脉硬化(0.278 9)、高脂血症(0.457 8)和糖尿病视网膜病变(0.314 3), 主要治疗药物包括门冬胰岛素注射液、胰岛素注射液和甲钴胺分散片等。结论: 2型糖尿病并发症群以周围神经病变、外周血管病变、脂代谢异常及微血管病变为主要特征。主题模型可从临床数据中提取出有临床意义的用药规律。

关键词 主题模型; 2型糖尿病; 治疗方案; 并发症; 实验室指标

Application of Topic Model in the Study of Type 2 Diabetes Treatment Plan

YANG Fan¹, LI Haiying¹, XU Shaojie¹, ZOU Zhaozhong¹, LI Yiqin², CHEN Huahong¹ (1. Dept. of Pharmacy, Zhongshan Bo'ai Hospital, Guangdong Zhongshan 528403, China; 2. Dept. of Endocrinology, Zhongshan Bo'ai Hospital, Guangdong Zhongshan 528403, China)

ABSTRACT OBJECTIVE: To explore the application of topic model in the study of type 2 diabetes treatment plan. METHODS: Clinical data of 650 inpatients with type 2 diabetes in our hospital during Jan. 2012-Jun. 2016 were analyzed retrospectively. The data of clinical diagnosis, lab indexes and clinical drug use were exchanged, summarized and merged by MATLAB R2014a software. Latent Dirichlet allocation and author topic model were adopted to extract the typical topics with topic

- 析[J]. 中国药业, 2015, 24(17): 69-70.
- [5] 钱军, 华海涛, 秦叔逵. 通关藤制剂抗肿瘤作用研究进展[J]. 中国中药杂志, 2009, 34(1): 11-13.
- [6] 国家药典委员会. 中华人民共和国药典: 四部[S]. 2015版. 北京: 中国医药科技出版社, 2015: 114-115.
- [7] 陈敏, 李媛媛, 李先荣. HPLC测定通关藤药材及消癌平片中通关藤苷A含量[J]. 中国中医药信息杂志, 2013, 20(3): 55-57.
- [8] 徐凯, 乔佳, 韩丽, 等. HPLC法测定不同产地通关藤中通关藤苷G和I的含量[J]. 药物分析杂志, 2016, 36(4): 607-610.
- [9] 李媛媛, 倪艳, 李先荣, 等. 通关藤药材C21甾体皂苷类成分特征图谱研究[J]. 中国实验方剂学杂志, 2011, 17(22): 43-46.
- [10] 方奕奇, 孙雪梅. 通关藤化学成分及抗肿瘤作用的研究进展[J]. 中国生化药物杂志, 2011, 32(2): 165-167.
- [11] 陶雪芬, 徐佳丽, 张如松. 萝藦科各属植物C21甾体皂苷类成分及药理活性研究进展[J]. 中国民族民间医药, 2009, 18(17): 40-42.
- [12] 刘玉芬, 孙凤英, 王玲. 乌骨藤化学成分研究[J]. 中草药, 1998, 20(9): 35-36.
- [13] 孟昭珂, 吴春霞. 通关藤新苷B的提取及体外抗肿瘤活性研究[J]. 海峡药学, 2009, 21(11): 34-37.
- [14] 邢旺兴, 唐婷, 谷娜, 等. 抗肿瘤中药通关藤的药理作用及其作用机理初步研究[J]. 健康研究, 2011, 31(1): 12-16.
- [15] 吴雪梅, 钟智. 中草药制剂与输液配伍的不溶性微粒考察[J]. 中国医院药学杂志, 1999, 19(2): 81-83.
- [16] 付文焕, 王斌, 施孝金. 注射剂中不溶性微粒相关研究现状及思考[J]. 上海医药, 2012, 33(21): 29-32.
- [17] 汪筠, 孙文武. 静脉输液不溶性微粒污染的实验研究[J]. 护理学杂志, 2006, 21(19): 72-74.
- [18] 王冰, 张梅英. 防范静脉输液中不溶性微粒危害的研究进展[J]. 上海护理, 2009, 9(1): 63-66.
- [19] 姜珊, 倪健. 中药注射液不溶性微粒的研究现状与再认识[J]. 中国药师, 2009, 12(10): 1465-1468.
- [20] 梁武英, 韩勇. 血栓通与3种溶媒配伍的稳定性观察[J]. 临床合理用药杂志, 2013, 6(7A): 44-45.

Δ 基金项目: 中山市科学事业费项目(医疗卫生)(No.2015B1180)

* 副主任药师。研究方向: 临床药学主题模型构建与数据挖掘。

电话: 0760-88776108. E-mail: 13928105684@139.com

(收稿日期: 2016-08-22 修回日期: 2017-05-23)

(编辑: 张元媛)

probability value > 0.1, and the topics was described by the complications with cumulative probability value > 0.5. RESULTS: A total of 62 complications words, 16 abnormal laboratory indexes groups and 20 treatment plans were obtained. A total of 4 typical topics were excavated (cumulative probability values for the first few complications were 0.878 6, 0.824 7, 0.821 5, 0.753 6; topic probability value were 0.336 4, 0.277 3, 0.203 5, 0.117 6, respectively) and were mainly characterized by peripheral neuropathy, peripheral vascular disease, abnormal lipid metabolism and microvascular lesions; abnormal lab indexes groups met the above characteristics. The complications with high distribution rate included diabetic peripheral neuropathy (0.578 7), hypertension (0.363 1), atherosclerosis (0.278 9), hyperlipidemia (0.457 8) and diabetic retinopathy (0.314 3); main drugs included Insulin aspart injection, Insulin injection, Methylcobalamin dispersible tablets, etc. CONCLUSIONS: The complications of type 2 diabetes are characterized by peripheral neuropathy, peripheral vascular disease, abnormal lipid metabolism and microvascular lesions. The medication rules with clinical significance can be extracted from the clinical data by topic model.

KEYWORDS Topic model; Type 2 diabetes; Treatment plan; Complications; Lab indexes

主题模型(Topic mode)能够提取隐含在文档(或其他离散数据集)中的主题,其中每个主题是语义相关词的多项式分布;主题模型的主要目的是提取数据集中隐含的统计规律,并利用主题进行直观表达,然后根据所获得的主题进行信息检索、分类、聚类、摘要提取以及信息间相似性、相关性判断等一系列应用^[1]。本研究将主题模型引入到临床用药规律的研究中,尝试应用该模型的无监督数据分析方法提取出有临床意义的用药规律,以期临床用药研究提供一种新颖的统计方法,为循证医学提供新的理论支持。

1 资料与方法

1.1 研究对象

收集2012年1月—2016年6月于我院住院治疗的2型糖尿病患者的临床数据。纳入标准:(1)符合《中国2型糖尿病防治指南(2013版)》中糖尿病诊断标准、糖代谢状态分类标准和分型体系^[2];(2)年龄≥18岁;(3)空腹血糖(FBG)≥7.8 mmol/L,糖化血红蛋白(HbA_{1c})≥7.5%;(4)临床资料完整。排除标准:(1)糖耐量受损(IGT)、空腹血糖异常(IFG)者;(2)确诊为1型糖尿病患者;(3)继发性糖尿病患者;(4)妊娠糖尿病患者;(5)合并严重的心、肝、肾、脑等疾病者。

本研究共纳入2型糖尿病患者650例。其中,男性347例(53.38%),女性303例(46.62%);年龄24~91岁,平均年龄(51.63±15.44)岁;首次住院368例(56.62%),再次住院282例(43.38%)。

1.2 研究方法

回顾性分析所有患者的临床数据,包括一般资料、临床诊断、相关实验室指标和临床用药情况等。主要内容如下:(1)患者的住院号、性别、年龄、出入院时间、住院次数;(2)出院诊断;(3)收缩压(SBP)、舒张压(DBP)、HbA_{1c}、FBG和餐后2h血糖(2hPBG)、空腹及餐后2h血清C肽(FCP、PCP);总胆固醇(TC)、三酰甘油(TG)、高密度脂蛋白胆固醇(HDL-C)、低密度脂蛋白胆固醇(LDL-C)、肌酐(Cr)、尿酸(UA)、天冬氨酸转氨酶(AST)、丙氨酸转氨酶(ALT);(4)药品名称、规格、剂量、给药方式、频率及医嘱起止时间等。

1.3 数据处理

采用MATLAB R2014a软件对患者的临床诊断、实验室指标和临床用药等数据进行转换、汇总及矩阵合并,共得到62个并发症词汇、16个异常实验室指标组和20个治疗方案。其中,并发症词汇包括以2型糖尿病为第一诊断且不重复的出院诊断词汇;异常实验室指标组是本研究汇总的异常实验室指标的组(见表1);治疗方案是从患者应用的108个药品的药理学汇总组合中筛选出的且应用频率>6%^[1]的20个治疗方案(见表2)。应用隐狄利克雷分配模型(Latent Dirichlet allocation, LDA)和作者主题模型(Author topic model, ATM)在三者的分配与扩展中进行数据挖掘,运行MATLAB R2014a软件中的“Statistics toolbox”模块,通过近似推理的方法获得隐变量(主题)的值,使用多个估算参数反复测量提取结果。以概率值>0.1的主题为典型主题,通过累计概率值>0.5的并发症对主题进行描述^[1]。

表1 16个异常实验室指标组

Tab 1 Sixteen abnormal lab indexes groups

编号	异常实验室指标组
1	FPG, 2 hPBG, HbA _{1c} , DBP, FCP, Cr, TC, LDL-C
2	FPG, 2 hPBG, HbA _{1c} , SBP, FCP, Cr, UA, TG, TC, HDL-C, LDL-C
3	FPG, 2 hPBG, HbA _{1c} , PCP, UA, ALT, AST, TG, TC
4	FPG, 2 hPBG, HbA _{1c} , UA, TG, TC, HDL-C
5	FPG, 2 hPBG, HbA _{1c} , SBP, DBP, UA, TG, TC, HDL-C
6	FPG, 2 hPBG, HbA _{1c} , ALT, TG, TC, LDL-C
7	FPG, 2 hPBG, HbA _{1c} , SBP, PCP, TG, HDL-C, LDL-C
8	FPG, 2 hPBG, HbA _{1c} , SBP, DBP, PCP, Cr, UA, TG, TC, HDL-C, LDL-C
9	FPG, 2 hPBG, HbA _{1c} , SBP, PCP, Cr, TG, HDL-C, LDL-C
10	FPG, 2 hPBG, HbA _{1c} , FCP, PCP, ALT, AST, TC, HDL-C
11	FPG, 2 hPBG, HbA _{1c} , SBP, DBP, Cr, ALT, TG, TC
12	FPG, 2 hPBG, HbA _{1c} , PCP, UA, TG, TC, HDL-C
13	FPG, 2 hPBG, HbA _{1c} , DBP, FCP, AST, TC, LDL-C
14	FPG, 2 hPBG, HbA _{1c} , SBP, DBP, PCP, Cr, TG, HDL-C, LDL-C
15	FPG, 2 hPBG, HbA _{1c} , SBP, Cr, ALT, TG, TC, HDL-C
16	FPG, 2 hPBG, HbA _{1c} , SBP, Cr, UA, TG, TC, HDL-C, LDL-C

2 结果

通过近似推理的方法把迭代次数设置为800次,将主题数目设定为8^[1],挖掘出主题概率值>0.1的4个典型主题,通过累计概率值>0.5的并发症对其进行描述,详见表3~6(各主题只列出并发症、治疗方案和异常实验室指标组等内容,主要描述了2型糖尿病的并发症:周

围神经病变、外周血管病变、脂代谢异常、微血管病变的分布概率、基本特征及治疗方案)。

表2 20个治疗方案

Tab 2 Twenty treatment plans

编号	治疗方案
1	短效胰岛素+长效胰岛素+ α -葡萄糖苷酶抑制剂+心脑血管疾病预防用药+调血脂药+抗高血压药
2	短效胰岛素+长效胰岛素+ α -葡萄糖苷酶抑制剂+周围神经病变用药+心脑血管疾病预防用药+调血脂药+抗高血压药
3	短效胰岛素+预混胰岛素+ α -葡萄糖苷酶抑制剂+周围神经病变用药+调血脂药
4	短效胰岛素+长效胰岛素+双胍类降糖药+周围神经病变用药+心脑血管疾病预防用药+抗高血压药
5	短效胰岛素+ α -葡萄糖苷酶抑制剂+周围神经病变用药+心脑血管疾病预防用药+调血脂药+抗高血压药
6	短效胰岛素+预混胰岛素+双胍类降糖药+ α -葡萄糖苷酶抑制剂+周围神经病变用药+调血脂药
7	短效胰岛素+预混胰岛素+心脑血管疾病预防用药+调血脂药+抗高血压药
8	短效胰岛素+长效胰岛素+预混胰岛素+双胍类降糖药+周围神经病变用药+心脑血管疾病预防用药+调血脂药+抗高血压药
9	短效胰岛素+长效胰岛素+ α -葡萄糖苷酶抑制剂+周围神经病变用药+心脑血管疾病预防用药+调血脂药
10	短效胰岛素+预混胰岛素+周围神经病变用药+心脑血管疾病预防用药+调血脂药+抗高血压药
11	短效胰岛素+周围神经病变用药+心脑血管疾病预防用药+调血脂药+抗高血压药
12	短效胰岛素+预混胰岛素+ α -葡萄糖苷酶抑制剂+周围神经病变用药+心脑血管疾病预防用药+调血脂药+抗高血压药
13	短效胰岛素+双胍类降糖药+ α -葡萄糖苷酶抑制剂+心脑血管疾病预防用药+调血脂药+抗高血压药
14	短效胰岛素+ α -葡萄糖苷酶抑制剂+心脑血管疾病预防用药+调血脂药+抗高血压药
15	短效胰岛素+双胍类降糖药+ α -葡萄糖苷酶抑制剂+周围神经病变用药+心脑血管疾病预防用药+调血脂药
16	短效胰岛素+长效胰岛素+预混胰岛素+心脑血管疾病预防用药+调血脂药+抗高血压药
17	短效胰岛素+双胍类降糖药+ α -葡萄糖苷酶抑制剂+心脑血管疾病预防用药+调血脂药
18	短效胰岛素+双胍类降糖药+ α -葡萄糖苷酶抑制剂+周围神经病变用药+心脑血管疾病预防用药+调血脂药+抗高血压药
19	短效胰岛素+长效胰岛素+双胍类降糖药+周围神经病变用药+心脑血管疾病预防用药+调血脂药+抗高血压药
20	短效胰岛素+长效胰岛素+预混胰岛素+周围神经病变用药+心脑血管疾病预防用药+调血脂药+抗高血压药

表3 主题1的并发症、治疗方案及异常实验室指标组概率分布

Tab 3 Probability distribution of complications, treatment plans and lab indexes group of topic 1

并发症	分布概率	治疗方案编号	分布概率	异常实验室指标组编号	分布概率
糖尿病周围神经病变	0.578 7	15	0.237 4	4	0.242 1
慢性浅表性胃炎	0.204 4	11	0.073 2	5	0.052 5
糖尿病足	0.063 7	2	0.156 1	2	0.047 3
心律失常	0.031 8	20	0.047 3	12	0.021 8

2.1 主题1结果分析

由表3可见,主题1的前几项并发症及其分布概率为:糖尿病周围神经病变(0.578 7)、慢性浅表性胃炎(0.204 4)、糖尿病足(0.063 7)和心律失常(0.031 8)。该

主题以周围神经病变为主要特征,且糖尿病周围神经病变的分布概率最高。

表4 主题2的并发症、治疗方案及异常实验室指标组概率分布

Tab 4 Probability distribution of complications, treatment plans and lab indexes group of topic 2

并发症	分布概率	治疗方案编号	分布概率	异常实验室指标组编号	分布概率
高血压	0.363 1	2	0.156 1	14	0.152 2
动脉硬化	0.278 9	18	0.122 7	7	0.123 5
脑梗死	0.094 4	20	0.063 8	15	0.074 3
高血压心脏病	0.046 9	7	0.041 6	14	0.136 5
冠心病	0.041 4	12	0.084 3	9	0.039 2

表5 主题3的并发症、治疗方案及异常实验室指标组概率分布

Tab 5 Probability distribution of complications, treatment plans and lab indexes group of topic 3

并发症	分布概率	治疗方案编号	分布概率	异常实验室指标组编号	分布概率
高脂血症	0.457 8	17	0.287 4	6	0.196 6
脂肪肝	0.246 5	17	0.287 4	3	0.098 6
电解质紊乱	0.094 9	9	0.067 2	15	0.074 3
低蛋白血症	0.022 3	20	0.037 5	11	0.015 0

表6 主题4的并发症、治疗方案及异常实验室指标组概率分布

Tab 6 Probability distribution of complications, treatment plans and lab indexes group of topic 4

并发症	分布概率	治疗方案编号	分布概率	异常实验室指标组编号	分布概率
糖尿病视网膜病变	0.314 3	5	0.156 7	16	0.088 3
糖尿病肾病	0.271 5	12	0.093 6	8	0.114 5
糖尿病足	0.167 8	2	0.156 1	2	0.047 3

该并发症群相对应的治疗方案及其分布概率为:治疗方案2(0.156 1)、治疗方案11(0.073 2)、治疗方案15(0.237 4)和治疗方案20(0.047 3)。该主题对应的治疗方案中应用频率较高的药物为门冬胰岛素注射液、胰岛素注射液、甲钴胺分散片、维生素B₆片、马来酸桂哌齐注射液、小牛血清去蛋白注射液和氟伐他汀钠缓释片。

该并发症群相对应的异常实验室指标组及其分布概率为:指标组2(0.047 3)、指标组4(0.242 1)、指标组5(0.052 5)和指标组12(0.021 8)。该主题异常实验室指标组显示,UA、TG、TC和HDL-C等指标异常,符合周围神经病变的基本特征^[3]。

该并发症群前4个并发症的累计概率值为0.878 6,该主题的概率值为0.336 4,且在所有主题中概率值最高,提示周围神经病变在2型糖尿病人群中占较高的比例。

2.2 主题2结果分析

由表4可见,主题2的前几项并发症及其分布概率为:高血压(0.363 1)、动脉硬化(0.278 9)、脑梗死(0.094 4)、高血压心脏病(0.046 9)和冠心病(0.041 4)。该主题以外周血管病变为主要特征,其中高血压和动脉硬化的分

布概率较高。

该并发症群相对应的治疗方案及其分布概率为:治疗方案2(0.156 1)、治疗方案7(0.041 6)、治疗方案12(0.084 3)、治疗方案18(0.122 7)和治疗方案20(0.063 8)。该主题对应的治疗方案中应用频率较高的药物为门冬胰岛素注射液、胰岛素注射液、甘精胰岛素注射液、马来酸桂哌齐特注射液、小牛血清去蛋白注射液、氟伐他汀钠缓释片、阿司匹林肠溶片、苯磺酸氨氯地平片、氯沙坦钾片和螺内酯片。

该并发症群相对应的异常实验室指标组及其分布概率为:指标组7(0.123 5)、指标组9(0.039 2)、指标组14(0.136 5或0.152 2)和指标组15(0.074 3)。该主题异常实验室生化指标组显示,SBP、TG、HDL-C和LDL-C等指标异常,主要表现为TG和LDL-C水平升高,而HDL-C水平降低,符合外周血管病变的基本特征^[4]。

该并发症群前5个并发症的累计概率值为0.824 7,该主题的概率值为0.277 3,提示外周血管病变在2型糖尿病病人中也占较高的比例。

2.3 主题3结果分析

由表5可见,主题3的前几项并发症及其分布概率为:高脂血症(0.457 8)、脂肪肝(0.246 5)、电解质紊乱(0.094 9)和低蛋白血症(0.022 3)。该主题主要以脂代谢异常为主要特征,其中高脂血症的分布概率较高。

该并发症群相对应的治疗方案及其分布概率为:治疗方案9(0.067 2);治疗方案17(0.287 4)和治疗方案20(0.037 5)。该主题对应的治疗方案中应用频率较高的药物为门冬胰岛素注射液、胰岛素注射液、氟伐他汀钠缓释片、阿托伐他汀钙片和盐酸二甲双胍片。

该并发症群相对应的异常实验室指标组及其分布概率为:指标组3(0.098 6)、指标6组(0.196 6)、指标组11(0.015 0)和指标组15(0.074 3)。该主题异常实验室指标组显示,TG、TC和ALT等指标异常,符合脂代谢异常的基本特征^[5-6]。

该并发症群前4个并发症的累计概率值为0.821 5,该主题的概率值为0.203 5,提示脂代谢异常在2型糖尿病病人中占一定比例。

2.4 主题4结果分析

由表6可见,主题4的前几项并发症及其分布概率为:糖尿病视网膜病变(0.314 3)、糖尿病肾病(0.271 5)和糖尿病足(0.167 8)。该主题主要以微细血管病变为主要特征,其中糖尿病视网膜病变与糖尿病肾病的分布概率较高。

该并发症群相对应的治疗方案及其分布概率为:治疗方案2(0.156 1)、治疗方案5(0.156 7)和治疗方案12(0.093 6)。该主题对应的治疗方案中应用频率较高的药物为门冬胰岛素注射液、胰岛素注射液、精蛋白生物合成人胰岛素注射液、伏格列波糖片、甲钴胺分散片、小牛血清去蛋白注射液、氟伐他汀钠缓释片、阿司匹林肠溶片和氯沙坦钾片。

该并发症群相对应的异常实验室指标组及其分布概率为:指标组2(0.047 3)、指标组8(0.114 5)和指标组16(0.088 3)。该主题异常实验室指标组显示,SBP、Cr、UA、TG、TC、HDL-C和LDL-C等指标异常,符合微细血管病变的基本特征^[7]。

该并发症群前3个并发症的累计概率值为0.753 6,该主题的概率值为0.117 6,提示微细血管病变在2型糖尿病病人中占一定比例。

3 讨论

主题模型的基本思想是假设存在 K 个(隐)主题,其中每个主题是实体间的多项式分布,即把每个主题描述为多个实体的概率分布;而每篇文档看作是这 K 个(隐)主题的一种随机混合,即把每篇文档描述为一些主题上的概率分布。主题模型的目标是利用主题结构来表现文档内或文档间的统计关系。给定一组文档,主题模型可以挖掘出这组文档讨论的核心主题以及每篇文档所涉及主题的分布概率。

本研究尝试应用“并发症-治疗方案-实验室指标”作为主题文档的低维表达形式,用于提取临床数据中的实验室指标、治疗方案和临床诊断或并发症的主题结构。该主题模型的核心思想是假设一类患者样本中包含多个并发症组合,每个并发症组合由一个主题表示,这些主题通过分析并发症和治疗方案获取,然后利用实验室指标对这些主题进行标记。其中,把每个主题看作是并发症上的多项式分布;同时,把每种治疗方案看作是主题上的多项式分布,因为一类治疗方案可以治疗多种并发症^[8];把实验室指标看作是主题上的多项式分布,因为同一组实验室指标可对多种并发症进行描述^[8]。

为了系统分析治疗过程中具有临床意义的多个实体的相关性,笔者利用该主题模型提取了临床数据中的并发症、治疗方案和实验室指标这3个实体的(隐)主题结构。该模型假设一类临床数据中包括多个主题,且这些主题所包含的信息特征具有相对的完整性和独立性,通过统计分析患者的实验室指标,找出其所隐含的主题结构,同时研究与之相对应的治疗方案以及每个主题的并发症描述特征,是一个由显变量(实验室指标)到隐变量(主题,通过并发症进行描述)再到显变量(治疗方案)的过程。

2型糖尿病是一种具有很多并发症的复杂疾病,有各种实验室指标和临床诊断结果。笔者利用统计主题模型对其并发症建模,并利用实验室指标、治疗方案特征确定主题内容,并发症则是对主题的描述。研究结果通过与2型糖尿病及其并发症的治疗进展、药物治疗研究进展、Meta分析等参考文献的结果和结论进行比对分析^[9-12],验证了本研究的结果符合循证医学规律,反映了以一种和/或多种并发症为主体的并发症群特征及治疗方案,从不同角度、不同层次分析和总结了2型糖尿病的并发症群与综合治疗方案^[13]。

主题模型及其在中医临床诊疗中的应用研究国内

多西他赛、表柔比星联合环磷酰胺对不同分子分型乳腺癌患者临床疗效及肿瘤标志物表达的影响

王松^{1*}, 薛文华¹, 赵晓玉¹, 梁艳¹, 齐跃东¹, 康健¹, 高明^{2#} (1. 郑州大学第一附属医院药学部, 郑州 450052; 2. 郑州大学第一附属医院肿瘤科, 郑州 450052)

中图分类号 R979.1 文献标志码 A 文章编号 1001-0408(2017)23-3212-05

DOI 10.6039/j.issn.1001-0408.2017.23.12

摘要 目的:探讨多西他赛、表柔比星联合环磷酰胺为新辅助化疗(NAC)方案对不同分子分型乳腺癌患者临床疗效及肿瘤标志物表达的影响。方法:选取2014年1月—2016年1月我院局部晚期乳腺癌女性患者88例,依据分子分型分为Luminal A型(23例)、Luminal B型(21例)、基底细胞型(11例)、HER2过表达型(18例)和正常乳腺样型(15例)。所有患者均予多西他赛注射液+盐酸表柔比星注射液+注射用环磷酰胺连续治疗6个周期(每21 d为1个周期),比较各分子分型乳腺癌患者的总有效率和病理完全缓解(pCR)率,以及治疗前后血清中肿瘤标志物[癌胚抗原(CEA)、糖蛋白(CA)125、CA153]的表达水平,记录不良反应的发生情况。结果:88例患者的总有效率为63.64%,其中基底细胞型乳腺癌患者的总有效率为72.73%,显著高于其他分子分型患者,差异均有统计学意义($P < 0.05$);而其他各分子分型患者总有效率两两比较,差异均无统计学意义($P > 0.05$)。88例患者的pCR率为27.27%,以基底细胞型乳腺癌患者的pCR率最高(45.45%);除HER2过表达型与正常乳腺样型患者pCR率无明显差异($P > 0.05$)外,其余各分子分型患者pCR率两两比较,差异均有统计学意义($P < 0.05$)。治疗前,各分子分型乳腺癌患者血清中CEA、CA125、CA153的表达水平比较,差异均无统计学意义($P > 0.05$);治疗后,各分子分型患者CEA、CA125、CA153的表达水平均较治疗前显著降低,差异均有统计学意义($P < 0.05$);但各分子分型患者上述标志物表达水平比较,差异均无统计学意义($P > 0.05$)。各分子

已有相关文献报道^[1,8],但尚未见其在非中医循证医学中的应用。主题模型是医工(理)结合的新型研究,但主题数目的设定、文档的低维表达、模型的扩展等内容仍在进一步的研究、探索过程中,以上问题也是本课题组后续研究的重点。本研究采用主题模型的方法对并发症、实验室指标和临床用药等大量临床数据进行数据挖掘,并通过并发症对主题进行描述,取得了研究预期的结果。结果表明,主题模型可以从临床数据中提取出有临床意义的用药规律,进一步从非中医循证医学的角度验证了主题模型能够为药物与临床研究提供一种新颖的统计方法和理论支持,为临床药理学基于知识与案例推理的决策支持方法提供了新的研究方向^[14-15]。但与国内已有的主题模型在中医领域的应用研究比较,纳入研究的病例数少、挖掘出来的典型主题个数较少是本研究的不足之处,有待后续研究进一步完善。

参考文献

- [1] 张小平. 主题模型及其在中医临床诊疗中的应用研究[D]. 北京:北京交通大学,2011.
- [2] 中华医学会糖尿病学分会. 中国2型糖尿病防治指南:2013版[J]. 中华糖尿病杂志,2014,6(7):447-498.
- [3] 杨秀颖,张莉,陈熙,等. 2型糖尿病周围神经病变机制研究进展[J]. 中国药理学通报,2016,32(5):598-602.
- [4] 念馨,李红. 2型糖尿病外周血管病变研究进展[J]. 医学

综述,2010,16(15):2341-2345.

- [5] 钟勇,江时森. 调脂药物对糖尿病状态下糖代谢的影响及其机制的研究进展[J]. 中华老年心脑血管病杂志,2013,15(9):990-992.
- [6] 张坤,任巧华,吴韬,等. 西格列汀联合二甲双胍治疗2型糖尿病合并代谢综合征的临床观察[J]. 中国药房,2016,27(17):2364-2366.
- [7] 杨立勇. 关注糖尿病微血管并发症研究进展[J]. 中华糖尿病杂志,2016,8(8):449-451.
- [8] 冯妍. 基于证候要素的不稳定心绞痛中西医结合治疗方案优化研究[D]. 北京:中国中医科学院,2012.
- [9] 马珊珊,张爱琴,姜晓璐,等. 2型糖尿病及其并发症的治疗进展[J]. 中国医药指南,2011,9(33):269-271.
- [10] 阳平. 糖尿病并发症的药物治疗研究进展[J]. 临床合理用药杂志,2010,3(21):153-155.
- [11] 王郡,周慧敏,董斌,等. 精蛋白生物合成人胰岛素注射液(预混30R)强化治疗初发2型糖尿病的临床研究[J]. 中国药房,2016,27(32):4522-4524.
- [12] 蔡晓凌,杨文嘉,高学营,等. 中国2型糖尿病患者胰岛素剂量的荟萃分析[J]. 中国糖尿病杂志,2016,24(6):490-507.
- [13] 霍蕊莉. 真实世界消渴病痹症证候特点与治疗方案研究[D]. 北京:中国中医科学院,2014.
- [14] 杨丽. 基于知识推理的中医临床诊疗决策支持方法研究[D]. 北京:北京交通大学,2014.
- [15] 毕澜馨. 基于案例推理的中医临床决策支持系统研究[D]. 北京:北京交通大学,2013.

* 主管药师,硕士。研究方向:药品调剂、临床药学。电话:0371-66862639。E-mail:470373072@qq.com

通信作者:副主任医师,博士。研究方向:肺癌、乳腺癌、胃癌等多种恶性肿瘤。电话:0371-66295542。E-mail:gaohaoming@163.com

(收稿日期:2016-10-30 修回日期:2017-03-28)

(编辑:张元媛)