

文章编号:1003-2754(2022)08-0741-05

doi:10.19845/j.cnki.zfysjbbz.2022.0185

# 基于 SMOTE 算法的脑血管介入术后 再狭窄风险预警模型的构建

雷雨, 左家财

**摘要:** **目的** 探索脑血管介入术后发生再狭窄风险的影响因素,并基于 SMOTE 算法建立预警模型。**方法** 选取绵阳市中心医院神经内科 2017 年 5 月 - 2021 年 5 月期间收治的 320 例脑血管狭窄患者,回顾性分析所有患者的病历资料,根据患者介入术治疗后是否发生再狭窄将其分为狭窄组和未狭窄组,采用单因素和多因素 Logistic 回归分析法筛选出脑血管介入术后发生再狭窄的风险因素并建立预测模型,同时基于 SMOTE 算法扩充阳性组数据,构建改进数据集的预警模型并对比验证模型的预测效能。**结果** 吸烟史、高脂血症、糖尿病、CRP  $\geq 5$  mg/L、支架长度  $\geq 16$  mm 是脑血管疾病患者介入治疗后发生再狭窄的独立风险因素 ( $P < 0.05$ )。基于上述风险因素建立预警模型  $P_1$  的 AUC = 0.872 (95% CI 0.821 ~ 0.923), 预警模型  $P_2$  的 AUC = 0.847 (95% CI 0.816 ~ 0.879), 两种预测模型的效能并无显著性差异且 AUC 均超过 0.75, 预测模型的效能较高。**结论** 基于吸烟史、高脂血症、糖尿病、CRP、支架长度原始数据和过抽样算法建立的预警模型均具有较高的预测性,医护人员可据此进行有效干预,进而减少脑血管疾病患者介入治疗后发生再狭窄的风险,改善患者预后,提高患者生存质量。

**关键词:** 脑血管疾病; 介入治疗; 风险因素; 回归方程; SMOTE 算法

中图分类号:R749.3

文献标识码:A

**Construction of restenosis risk warning model based on SMOTE algorithm after cerebrovascular intervention** LEI Yu, ZUO Jiakai. (Mianyang Central Hospital, School of Medicine, University of Electronic Science and Technology of China, Mianyang 621000, China)

**Abstract:** **Objective** To explore the influencing factors of restenosis risk after cerebral vascular intervention, and establish a warning model based on SMOTE algorithm. **Methods** A total of 320 patients with cerebrovascular stenosis admitted to the Department of Neurology of Mianyang Central Hospital from May 2017 to May 2021 were selected. The medical records of all patients were retrospectively analyzed. According to whether restenosis occurred after interventional treatment, the patients were divided into stenosis group and non-stenosis group. Single factor and multi-factor Logistic regression analysis were used to screen the risk factors of restenosis after cerebrovascular intervention and establish the prediction model. At the same time, expand the positive group data based on SMOTE algorithm, build the early warning model of improved data set and compare and verify the prediction efficiency of the model. **Results** Smoking history, hyperlipidemia, diabetes mellitus, CRP  $\geq 5$  mg/L and stent length  $\geq 16$  mm were independent risk factors for restenosis after intervention therapy in patients with cerebrovascular diseases ( $P < 0.05$ ). Based on the above risk factors, the AUC of early warning model  $P_1$  and  $P_2$  was established to be 0.872 (95% CI 0.821 ~ 0.923) and 0.847 (95% CI 0.816 ~ 0.879), respectively. There was no significant difference in the efficacy of the two prediction models and their AUC was over 0.75, indicating that the prediction model had high efficacy. **Conclusion** Based on the history of smoking, hyperlipidemia, diabetes, CRP, bracket and the length of the original data of sampling algorithm to establish early warning model has higher predictive, medical personnel will enable effective intervention, thus reduce the patients with cerebrovascular disease risk for restenosis after interventional treatment, improve the prognosis of patients, improve the quality of survival.

**Key words:** Cerebrovascular disease; Interventional therapy; Risk factors; Regression equation; SMOTE algorithm

脑血管疾病是一类因脑血管病变或血流障碍导致脑功能障碍的疾病总称,主要包括颅内、外血管的闭塞狭窄、受损及破裂而引发的局限性或弥散性脑功能受损<sup>[1]</sup>。流行病学调查显示,我国每年新增脑血管疾病患者病例约 240 万,而死亡病例可达到 110 万,是危害中老年人群身体生命健康的最主要

疾病之一<sup>[2]</sup>。脑卒中是一种急性脑血管疾病,其快

收稿日期:2022-04-19; 修订日期:2022-05-30

作者单位:(四川省电子科技大学医学院附属绵阳医院-四川省绵阳市中心医院神经内科,四川 绵阳 621000)

通讯作者:左家财, E-mail: zuojiacai0409@163.com

速发作的特征使其成为患者残疾甚至死亡的全球第二大致死原因<sup>[3]</sup>。根据发病机制的不同,脑卒中可分为出血性卒中和缺血性卒中,前者主要是由于脑血管破裂而引起的神经功能缺损,而缺血性卒中则是由血管狭窄闭塞引起的脑组织缺氧损害<sup>[4]</sup>。颅内动脉粥样硬化性动脉狭窄是缺血性卒中的重要原因,在人群中约5%~10%的脑卒中是由粥样硬化性动脉狭窄引起<sup>[5]</sup>。有报道显示,与其他卒中亚型相比,由动脉狭窄引起的卒中与复发性短暂性脑缺血的发展有着较高的相关性<sup>[6]</sup>。目前临床上对于脑血管狭窄患者的最佳治疗选择是采用颅内动脉支架置入术。近年来,随着导管技术的发展和成熟,越来越多地患者在接受内科药物治疗的同时,选择进行脑血管介入术的治疗并取得了良好的治疗效果。然而,部分患者治疗后继续遭受脑血管再狭窄而导致复发性缺血事件的发生。为进一步预防和改善此类问题的发生,本研究通过回顾我院320例患者的病例资料,试图找出引起其发生再狭窄的风险因素,现报道如下。

## 1 资料与方法

### 1.1 研究对象

选取绵阳市中心医院神经内科2017年5月-2021年5月期间收治的320例脑血管狭窄患者,回顾性分析所有患者的病历资料,根据患者行介入术治疗后是否发生再狭窄将其分为狭窄组和未狭窄组,观察对比两组患者的一般资料,包括患者的性别、年龄、BMI、吸烟史、饮酒史、高脂血症、高血压、糖尿病、冠心病、血管类型(颈内动脉、椎动脉、颅内动脉)、心率、CRP、淋巴细胞计数、中性粒细胞计数、血小板计数、支架长度、支架直径等。

纳入标准:(1)经头部MRI检查证实为脑血管狭窄患者,符合颈内动脉、椎动脉以及颅内动脉狭窄的临床特征;(2)于我院行支架介入术治疗且无脑血管介入治疗史;(3)术后规律进行二级预防措施;(4)自愿参加,配合医护并签署知情同意书。排除标准:(1)患有感染性疾病患者;(2)合并肝肾功能异常及各类型的严重肿瘤患者;(3)近期存在颅脑外伤、颅内出血等疾病患者;(4)不配合医护治疗、病历资料不完善者及精神病患者等。

### 1.2 SMOTE 过抽样算法

SMOTE 过抽样算法:SMOTE 过抽样算法是将少量样本通过一定步骤(参照 Blagus<sup>[8]</sup>等的论文)扩充到一定量的一种方法,本研究中,少数量样本为再狭窄组=28例,需扩充倍数 $n = \text{未再狭窄组} / \text{再狭窄组} = 10.429$ ,取10倍数,为 $28 \times 10 = 280$ 例。通过此方法扩充的数据集,未改变样本的原有空间边界,具有较高的信效度。

### 1.3 统计学方法

数据均采用SPSS软件进行整理统计,其中计数资料以例数表示,计量资料以均值±标准差( $\bar{x} \pm s$ )表示,分别采用 $\chi^2$ 检验与 $t$ 检验, $P < 0.05$ 表示数据差异显著。采用 Logistic 回归分析筛选独立风险因素,SPSS 软件制作 ROC 曲线进行预测价值验证,以  $AUC > 0.75$  为预测价值较高。

## 2 结果

### 2.1 两组基本信息资料单因素分析

两组患者的年龄、BMI、吸烟史、高脂血症、高血压、糖尿病、冠心病、颅内动脉、CRP、中性粒细胞计数及支架长度的数据差异有统计学意义( $P < 0.05$ )(见表1)。

### 2.2 多因素 Logistic 回归分析结果

对2.1中具有显著性差异的因素作二元 Logistic 回归分析,结果表明:吸烟史、高脂血症、高血压、糖尿病、CRP、及支架长度为脑卒中患者行介入术后再狭窄的独立风险因素,且数据差异具有统计学意义( $P < 0.05$ )(见表2)。基于上述各独立风险因素及其回归系数,我们建立原始预警模型  $P_1 = -6.507 + 1.527 \times \text{吸烟史} + 1.544 \times \text{高脂血症} + 1.598 \times \text{糖尿病} + 1.145 \times \text{CRP} + 1.700 \times \text{支架长度}$ ,对该模型进行 Hosmer-Lemeshow 检验,结果可见:决定性系数  $R^2 = 0.340$ ,  $P = 0.795$ ,表明模型拟合度良好。

### 2.3 基于 SMOTE 过抽样算法的 logistic 预警模型

基于2.3中筛选出的独立危险因素,我们利用 SMOTE 过抽样算法进行10倍过抽样,使得术后再狭窄患者数和术后未再狭窄患者数比例近似为1(过抽样为未再狭窄患者=292例,再狭窄患者280例,比例 $\approx 1.043$ ),然后对过抽样后的数据重新拟合 Logistic 回归模型,结果(见表3)。基于 SMOTE 过抽样算法的预警模型  $P_2 = -3.583 + 1.159 \times \text{吸烟史} + 1.335 \times \text{高脂血症} + 1.293 \times \text{糖尿病} + 1.265 \times \text{CRP} + 1.454 \times \text{支架长度}$ 。同样对该模型进行 Hosmer-Lemeshow 检验,结果可见:决定性系数  $R^2 = 0.346$ ,  $P = 0.881$ ,表明模型拟合度良好。

### 2.4 预测模型的评估

为验证预测模型的效能及对比两组模型的优异,分别对预警模型  $P_1$  和  $P_2$  进行 ROC 曲线分析,结果可见,预警模型  $P_1$  的 ROC 曲线下面积 = 0.872 (95% CI 0.821 ~ 0.923), 预警模型  $P_2$  的 ROC 曲线下面积 = 0.847 (95% CI 0.816 ~ 0.879) (见图1、图2)。可见基于 SMOTE 过抽样算法扩充样本达到平衡后,预测模型的效能与原样本模型的效能并无差异,两组 AUC 均超过 0.75,预测模型的效能较高。

表1 两组基本信息资料单因素分析

分组	性别(男/女)	年龄(岁)	BMI(kg/m <sup>2</sup> )	吸烟史 n(%)	饮酒史 n(%)	高脂血症 n(%)	高血压 n(%)
狭窄组(n=28)	19/9	56.25 ± 15.30	26.82 ± 3.08	24(85.71)	10(35.71)	18(64.29)	25(89.29)
未狭窄组(n=292)	182/110	50.44 ± 13.74	23.36 ± 2.74	172(58.90)	88(30.14)	89(30.48)	196(67.12)
$\chi^2/t$	0.334	2.116	6.313	7.738	0.374	13.12	5.874
P值	0.563	0.035	<0.001	0.005	0.301	<0.001	0.015

分组	糖尿病 n(%)	冠心病 n(%)	颈内动脉 n(%)	椎动脉 n(%)	颅内动脉 n(%)	心率(次/分)	CRP(mg/L)
狭窄组(n=28)	21(75.00)	5(17.86)	10(35.71)	13(46.43)	5(17.86)	75.25 ± 8.66	6.85 ± 1.04
未狭窄组(n=292)	122(41.78)	21(7.19)	130(44.52)	141(48.29)	21(7.19)	77.08 ± 9.32	3.87 ± 1.16
$\chi^2/t$	11.407	3.893	0.805	0.035	3.893	0.998	13.095
P值	0.001	0.048	0.37	0.851	0.048	0.319	<0.001

分组	淋巴细胞计数(×10 <sup>9</sup> /L)	中性粒细胞计数(×10 <sup>9</sup> /L)	血小板计数(×10 <sup>9</sup> /L)	支架长度(mm)	支架直径(mm)
狭窄组(n=28)	5.22 ± 0.95	2.51 ± 1.06	274.24 ± 66.34	19.42 ± 6.33	3.85 ± 1.05
未狭窄组(n=292)	5.38 ± 1.21	1.71 ± 0.59	268.18 ± 60.30	14.58 ± 4.74	4.01 ± 1.35
$\chi^2/t$	0.68	6.285	0.504	4.98	0.609
P值	0.498	<0.001	0.615	<0.001	0.543

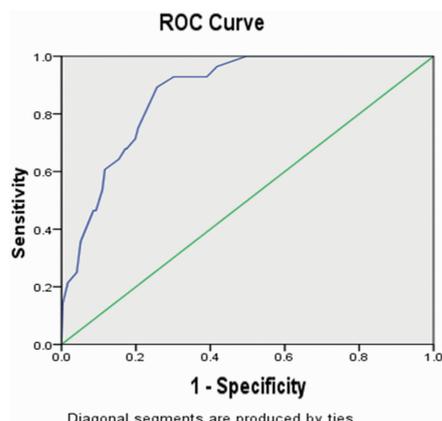
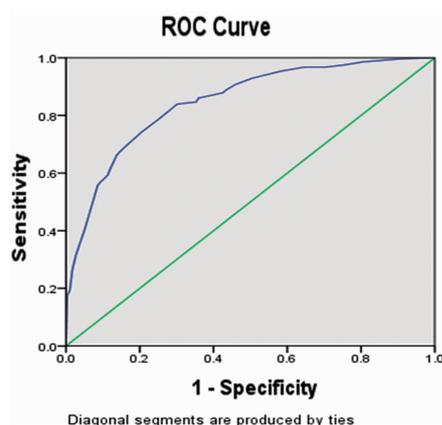
表2 多因素 Logistic 回归分析结果

因素	回归系数	标准误	Wald 值	P 值	OR	95% 置信区间	
						上限	下限
吸烟史	1.527	0.613	6.204	0.013	4.602	1.385	15.299
高脂血症	1.544	0.467	10.955	0.001	4.684	1.877	11.686
糖尿病	1.598	0.504	10.053	0.002	4.945	1.841	13.284
CRP	1.145	0.476	5.796	0.016	3.143	1.237	7.985
支架长度	1.7	0.469	13.13	<0.001	5.474	2.183	13.73
Constant	-6.507	0.913	50.821	<0.001	0.001		

注:自变量赋值,吸烟史(是=1,否=0);高脂血症(是=1,否=0);糖尿病(是=1,否=0);CRP(≥5 mg/L=1, <5 mg/L=0);支架长度(≥16 mm=1, <16 mm=0)

表3 基于 SMOTE 过抽样算法的 Logistic 回归分析

因素	回归系数	标准误	Wald 值	P 值	OR	95% 置信区间	
						上限	下限
吸烟史	1.159	0.248	21.896	<0.001	3.186	1.961	5.177
高脂血症	1.335	0.214	38.856	<0.001	3.8	2.497	5.782
糖尿病	1.293	0.217	35.663	<0.001	3.645	2.384	5.572
CRP	1.265	0.214	35.036	<0.001	3.543	2.331	5.387
支架长度	1.454	0.22	43.887	<0.001	4.282	2.785	6.584
Constant	-3.583	0.328	119.353	<0.001	0.028		

图1 预警模型 P<sub>1</sub> 的 ROC 曲线图2 预警模型 P<sub>2</sub> 的 ROC 曲线

### 3 讨论

随着社会生活方式的不断改变,心脑血管疾病的发生率不断增加,脑血管疾病现已列入我国致残率首位的疾病<sup>[7]</sup>。缺血性卒中是脑血管疾病的一种,其不仅影响了患者的语言、吞咽、肢体等功能,同时也在很大程度上降低了患者的日常生活质量<sup>[8]</sup>。血管内介入治疗的应用,对于药物治疗效果欠佳及无法耐受外科手术的患者有着极为重要的意义,正逐渐成为脑血管狭窄的重要治疗方法<sup>[9]</sup>。然而,血管内介入治疗同样有着较为明显的劣势,患者在治疗过程中可能出现血管损伤、破裂等并发症,术后也存在部分患者出现再狭窄风险<sup>[10]</sup>。有报道指出,血管内皮细胞由于支架置入时被损伤,受损的动脉内膜在自我修复时,由于支架的长期压迫而导致内膜组织增生过度,最终导致再狭窄的发生<sup>[11]</sup>。当前,脑血管介入治后再狭窄已成为脑血管疾病患者和医护不容忽视的一个难题,因此探寻相关风险因素,针对性的进行预防治疗,对脑血管疾病患者的预后有着重要的意义。

本研究通过回顾性分析,采用单因素和多因素 Logistic 回归分析法筛选出吸烟史、高脂血症、糖尿病、CRP $\geq 5$  mg/L、支架长度 $\geq 16$  mm 是脑血管疾病患者介入治疗后发生再狭窄的独立风险因素( $P < 0.05$ )。多项研究表明,烟草中的各类有害物质是造成心脑血管疾病的发病原因之一,其中尼古丁可引发氧化应激反应和炎症反应,对内皮细胞的功能有着严重的危害<sup>[12,13]</sup>。此外,吸烟可刺激患者的交感神经系统,促使血管不断收缩和扩张,致使脑血管的弹性下降,出现再狭窄的风险增大<sup>[14]</sup>。从动物实验中证实,尼古丁可导致各类炎性体活化及产生大量白细胞介素-1、活性氧等物质,加剧内皮细胞的炎症致其死亡<sup>[15]</sup>。本研究中,吸烟的脑血管疾病患者发生再狭窄风险的几率是不吸烟患者的 4 倍有余,验证了以上观点。高脂血症患者的脂质代谢异常,同样作为再狭窄的风险因素常被学者们考虑和研究<sup>[16]</sup>。有研究显示,脑血管介入术治疗时,支架的置入会导致血管内皮细胞损害,使得内皮覆盖不充分,从而允许更多的脂蛋白进入内膜下空间,最终形成富含脂质的泡沫纸巨噬细胞聚集而导致再狭窄的发生<sup>[17]</sup>。对于糖尿病的研究中发现,体内持续的高血糖对内皮细胞的损害巨大,其可以不断促进血管壁胶原的非酶糖基化从而加速血管硬化,还可以提高血液浓度,增加血栓发生的几率<sup>[18]</sup>。脑血管疾病患者在病灶处置入支架后,粘稠的血液极易粘附在支架上形成新的闭塞,慢慢累积下便会再狭窄的发生<sup>[19]</sup>。本研究中,再发生狭窄组患者中有 75% 以上的患者伴有糖尿病,显著高于未狭窄组的 41.78%。CRP 是一种极为敏感的炎症性标志物,常被用来判断各类炎症性疾病的发生。本研究发现,与未再狭窄组的(3.87 $\pm$ 1.16)mg/L 相比,发生再狭窄组患者的 CRP 水平达到(6.85 $\pm$ 1.04)mg/L,两组具有显著性差异。有研究<sup>[20]</sup>表明,血管损伤及血管内膜的增生均与炎性反应的发生有关,因此,CRP 也可纳入风险因素的考虑之中。脑血管介入治疗最主要的便是支架的置入,因此支架的长度、直径均可能对治疗效果及预后产生一定的影响。我们发现较长的支架发生再狭窄的风险几率较高,究其原因可能在于较长的支架会对血管内皮细胞的受损较大,从而回到上述几项观点中。

当下,传统的基于样本量平衡的各种统计学方

法已不适用医学领域类的不平衡数据集。有研究<sup>[21]</sup>显示,多数医学领域中,阳性样本的数量显著低于阴性样本数量,存在高度的不平衡状态,这就导致预测结果的特异度与灵敏度较差,最终对预警模型的预测精准度产生巨大的影响。SMOTE 过抽样算法是一种针对不平衡数据集的有效处理方法。因此,我们对上述几项风险因素进行了回归分析并建立了预警模型的同时,基于 SMOTE 过抽样算法,有效的扩充了阳性样本量,即再狭窄组例数,产生新的平衡数据集并再次建立预警模型。通过对两种预测模型的验证和对比,基于上述风险因素建立预警模型  $P_1$  的 ROC 曲线下面积 = 0.872 (95% CI 0.821 ~ 0.923), 基于 SMOTE 算法改进数据集的预警模型  $P_2$  的 ROC 曲线下面积 = 0.847 (95% CI 0.816 ~ 0.879), 可见基于 SMOTE 过抽样算法扩充样本达到平衡后,预测模型的效能与原样本模型的效能并无差异,两组 AUC 均超过 0.75,预测模型的效能较高。

综上所述,基于吸烟史、高脂血症、糖尿病、CRP、支架长度过抽样算法建立的预警模型具有较高的预测性,医护人员可以根据该模型在早期进行有效的防范与干预,进而减少脑血管疾病患者介入治疗后发生再狭窄的风险,改善患者预后,提高患者生存质量。

### [参考文献]

- [1] 刘 鸣,刘峻峰,吴 波. 脑血管病分类分型进展与解读[J]. 中华神经科杂志,2017,50(3):163-167.
- [2] 孙海欣,王文志. 中国 60 万人群脑血管病流行病学抽样调查报告[J]. 中国现代神经疾病杂志,2018,18(2):83-88.
- [3] Doria JW, Forgaes PB. Incidence, Implications, and Management of seizures following ischemic and hemorrhagic stroke[J]. Curr Neurol Neurosci Rep,2019,19(7):1-8.
- [4] Barthels D, Das H. Current advances in ischemic stroke research and therapies[J]. Biochim Biophys Acta Mol Basis Dis,2020,1866(4):1-10.
- [5] 张 磊,洪 波. 颅内动脉粥样硬化性狭窄介入治疗的问题和挑战[J]. 中国脑血管病杂志,2020,17(3):113-116.
- [6] Yaghi S, de Havenon A, Rostanski S, et al. Carotid stenosis and recurrent ischemic stroke: A post-hoc analysis of the POINT trial [J]. Stroke,2021,52(7):2414-2417.
- [7] 袁俊亮,李讓婷,王双坤,等. 脑小血管病研究进展[J]. 中华老年医学杂志,2020,39(4):481-485.
- [8] 梁宏艳,赵文博,马红蕊,等. 急性缺血性卒中取栓治疗延误的影响因素分析[J]. 中国脑血管病杂志,2020,17(2):57-62.
- [9] 朱旭成,陈荣华,邵华明,等. 症状性慢性大脑中动脉闭塞血管内再通治疗的初步分析[J]. 中国脑血管病杂志,2020,17(11):677-682.
- [10] 邓一鸣,张净瑜,高峰,等. 高龄老年急性缺血性脑卒中患者桥接治疗与直接介入治疗的对比分析[J]. 中华老年医学杂志,2020,39(10):1137-1141.
- [11] 席亚东,黄玉华,杜若林,等. 血管内支架植入后的内皮损伤及其修复策略[J]. 生物医学工程学杂志,2018,35(2):307-312.
- [12] Xue R, Wan Y, Sun X, et al. Nicotinic mitigation of neuroinflammation and oxidative stress after chronic sleep deprivation[J]. Front Immunol,2019,10:2546.
- [13] 涂洪波,陈刘通,孙 丽,等. 二甲双胍对尼古丁诱导肺动脉平滑肌细胞增殖的影响及机制研究[J]. 第三军医大学学报,2015,37(7):623-628.
- [14] 邹武松,余意君,王志红,等. 社区高血压患者自主神经功能检测及其价值研究[J]. 中国全科医学,2021,24(1):65-69.
- [15] 窦浚峰,海国栋,封青川,等. 尼古丁对白细胞介素-1 $\beta$  诱导的软骨细胞炎症反应及核因子- $\kappa$ B 信号通路的影响[J]. 中华实验外科杂志,2018,35(5):885-887.
- [16] Sun X, Feng M, Lu L, et al. Lipid abnormalities in patients with Cushing's disease and its relationship with impaired glucose metabolism [J]. Front Endocrinol (Lausanne),2021,11(32):1-9.
- [17] Nitzsche A, Poittevin M, Benarab A, et al. Endothelial SIP<sub>1</sub> signaling counteracts infarct expansion in ischemic stroke[J]. Circ Res,2021,128(3):363-382.
- [18] Chen Y, Hua Y, Li X, et al. Distinct types of cell death and the implication in diabetic cardiomyopathy [J]. Front Pharmacol,2020,11(18):1-9.
- [19] Wohrle J, Scheller B, Seeger J, et al. Impact of diabetes on outcome with drug-coated balloons versus drug-eluting stents: The BASKET-SMALL 2 trial [J]. JACC Cardiovasc Interv,2021,14(16):1789-1798.
- [20] Giacinto O, Satriano U, Nenna A, et al. Inflammatory response and endothelial dysfunction following cardiopulmonary bypass: Pathophysiology and pharmacological targets [J]. Recent Pat Inflamm Allergy Drug Discov,2019,13(2):158-173.
- [21] 顾永春,顾兴全,武 娇,等. 面向不平衡网络评论数据挖掘的服务质量评价[J]. 小型微型计算机系统,2021,42(2):354-361.