

费用数据离群值裁剪方法对某市 CHS-DRG 分组效能的影响*

北京大学公共卫生学院卫生政策与管理系(100191) 张靓因 李璟媛 冯文[△]

【摘要】 目的 支付标准的测算是 DRG 实施中的关键环节,本研究旨在分析比较四种常用的数据裁剪方法,为实际付费做好准备,也为后续的研究与实践提供启示及方法改进的依据。方法 利用中段区间法和缩尾法对某市二级医院 2017-2019 年出院病例的住院费用离群值进行裁剪,用变异系数、总体方差减少系数、Kruskal-Wallis H 统计量及 ROC 曲线评价不同方法裁剪后对 DRG 分组效能的影响。结果 四种方法裁剪离群值后对 DRG 分组效能指标均有改善,但改善程度各有不同,最适宜的裁剪方法为方法 2(裁剪上限 $Q_3 + 1.5IQR$, 裁剪下限 $Q_1 - 0.5IQR$)。综合评价离群值裁剪前后 CV、RIV 以及 ROC 曲线的变化,可以得到实现更佳分组效能的数据裁剪方法。结论 恰当裁剪离群值可规避医保基金的不必要浪费、维持医院经济平稳运行、减少患者疾病负担,地方医保局可根据实际住院数据恰当选取裁剪方法。

【关键词】 疾病诊断相关分组 离群值 数据裁剪

【中图分类号】 R197.1

【文献标识码】 A

DOI 10.3969/j.issn.1002-3674.2022.03.006

The Influence of Outlier Trimming Methods of Hospitalization Expenses on Efficiency of CHS-DRG

Zhang Liangnan, Li Jingyuan, Feng Wen (Department of Health Policy and Management, Peking University School of Public Health(100191), Beijing)

【Abstract】 Objective The calculation of payment standards is the key link of the implementation of DRG. This study aimed to analyze and compare the characteristics of the four commonly used data trimming methods, prepare for actual payment and provide enlightenment and method improvement basis for subsequent researches in related fields. **Methods** The outliers of hospitalization expenses of discharged cases in a city's secondary hospitals from 2017 to 2019 were trimmed using the interquartile range and winsorize. Using the coefficient of variation, reduction in variance (RIV), Kruskal-Wallis H and ROC curve to evaluate the improvement of DRG grouping performance by different methods. **Results** The four methods all improved the DRG grouping performance index after trimming the outliers, but the degree of improvement was different. The author believed that the most appropriate trimming method was method 2 (high trim point: $Q_3 + 1.5IQR$, low trim point: $Q_1 - 0.5IQR$). By comprehensively evaluating the changes of CV, RIV and ROC curves before and after outlier trimming, a data trimming method with better grouping performance could be obtained. The results of this study showed that the second method of trimming data could get better effectiveness. **Conclusion** Appropriate trimming could avoid unnecessary waste of medical insurance funds, maintain the stable operation of the hospital's economy and reduce the burden of patients' disease. The local medical insurance bureau should appropriately select trimming method according to the actual hospitalization data.

【Key words】 Diagnosis related groups; Outliers; Trimming data

疾病诊断相关分组 (diagnosis related groups, DRG) 是综合考虑患者的疾病诊断、手术操作、治疗方式、个体特征等因素,将病例进行分类的一种患者分组方案^[1]。该方法关注的是“临床过程”和“资源消耗”两个维度,分组结果要保障同一个 DRG 内的病例临床过程相似,资源消耗相近^[2],进而以 DRG 组为单位制定费用标准进行付费。

在根据历史出院病例进行 DRG 分组及病组均数计算、权重测算时,若组内包含远高于大多数患者住院费用水平的病例,将会导致该组的住院费用平均值被高估^[3],使该组的付费权重测算值增大,不仅会占用其余病组的预付金额,还会带来医保基金的浪费。反之,该病组的平均费用将被低估,付费权重测算值偏低,会使该类病组收不抵支,可能会导致医院推诿该类

患者,不仅会对医院管理决策产生重大影响,也将给患者带来沉重经济负担^[4]。

2019 年 10 月《国家医疗保障疾病诊断相关分组 (CHS-DRG) 分组与付费技术规范》中提出“精确付费是 DRG 实施的重要保障”。同年,国家医保局、财政部、国家卫生健康委、国家中医药局联合印发的《关于印发按疾病诊断相关分组付费国家试点城市名单的通知》中指出“对于付费异常高值或异常低值的病组,可按项目付费”,但离群值裁剪的指标及范围如何界定目前尚无统一规定。

中段区间法(也称四分位间距法)(interquartile range, IQR)和缩尾法(winsorize)是目前最为常见的患者住院费用离群值裁剪的两种方法。在不同的研究中,研究者又设置了不同的参数。如欧洲学者分别在 2003^[3]和 2008 年^[5]的研究中采用了中段区间法,但后者对极低界值的设定参数与前者不同。在缩尾法中,

* 基金项目:泰康溢彩公共卫生及流行病防治基金(No. 7)

[△]通信作者:冯文, E-mail: fengwenmail@sina.com

取上下 2.5%^[6] 和上下 1%^[7-8] 是常见的临界值设定方法。在国内各地试行 DRG 之后,学者在不同的研究中选择了不同的数据裁剪方法,如张宏等人的研究^[9] 选用住院费用的第 75 百分位数加 1.5 倍四分位数间距的方法裁剪高离群值,袁丽萍等人的研究^[10] 剔除了住院费用小于 1% 或大于 99% 的病例等。国际上,Francesc Cots^[3] 等人的研究中比较了不同离群值裁剪方法的效果和影响,而国内还鲜见相关比较研究。

本研究旨在通过对某市二级医院 2017-2019 年出院病例的住院费用进行裁剪,分析比较四种常用的数据裁剪方法的影响,从而选取较为合适的方法,达到降低 DRG 病组内数据变异程度、提升分组效能、提高 DRG 病组权重测算可靠性的目的,为后续实际付费做好准备,同时也为后续相关领域的研究提供启示及方法改进的依据。

资料与方法

1. 资料来源

本研究采用 2020 年 6 月 18 日国家医保局印发的《医疗保障疾病诊断相关分组(CHS-DRG)细分组方案(1.0 版)》(共 618 组)对我国西南地区某市 21 家二级医院 2017-2019 年 52.8 万条出院病例进行疾病分组,以 DRG 病组例数 ≥ 30 例为稳定病组,保留 526739 条出院病例,共计 369 组。

2. 方法

(1) 数据裁剪方法

本研究采用四种常用的离群值裁剪方法,具体公式如下:

方法 1: 裁剪上限 = $Q3 + 1.5 \times IQR$, 裁剪下限 = $Q1 - 1.5 \times IQR$, 式中, IQR 是住院费用分布的下四分位数($Q1$)和上四分位数($Q3$)之间的差值^[3,11]。

方法 2: 裁剪上限 = $Q3 + 1.5 \times IQR$, 裁剪下限 = $Q1 - 0.5 \times IQR$ ^[5,12]。

方法 3: 针对 $CV > 1$ 的病组,按照方法 1 公式裁剪(为避免过度裁剪,该方法只裁剪存在较大变异问题的病组)。

方法 4: 裁剪上限 = 住院费用的第 97.5 个百分位数,裁剪下限 = 住院费用的第 2.5 个百分位数^[13-14]。

(2) 分组效能评价指标

① 组内异质性评价: 变异系数(coefficient of variation, CV) 是检验统计数据聚类拟合度的重要指标^[15], 国际上通常把变异系数作为从 ADRG 到 DRG 寻找分类节点的标志^[16]。本研究用 DRG 组内住院费用的变异系数,反映各 DRG 组内不同病例住院费用的差异度(离散度),原则上要求组内变异系数小于 1。CV 值越小,说明组内的变异越小,分组效果越好^[10]。

$$CV = \frac{\text{DRG 组内住院费用标准差}}{\text{DRG 组内所有住院病例的次均费用}}$$

② 组间异质性评价

总体方差减少系数(reduction in variance, RIV): 为 DRG 组间差异度, CHS-DRG 用 RIV 来衡量 DRG 系统的分组效能, RIV 越大说明 DRG 组间异质程度越强, 分组系统的区分度越高, 分组效果越好。通常要求 RIV 值达到 70% 以上^[17]。

$$RIV = \frac{\text{总离均差平方和} - n \text{ 个子集住院费用的离均差平方和}}{\text{总离均差平方和}}$$

完全随机设计秩和(Kruskal-Wallis H)检验: 可检验不符合正态分布的多个独立样本间是否来自不同的总体分布, 当 $P < \alpha$ (α 为检验水准) 差异有统计学意义时, 认为多个总体间相比较有差异^[18-19]。本研究中, 各病组的次均住院费用不满足正态分布, 故使用 KW 统计量检验各病组经不同裁剪方法得到的各组次均费用是否有统计学差异。

③ 分组效果合理性的综合评价: 受试者工作曲线(receiver operator characteristic curve, ROC) 的检验变量是分别用方法 1、方法 2、方法 4 判断病例是否为极值, 状态变量是经裁剪强度最小的方法 3 判断病例是否为极值。

(3) 统计学方法

采用 Excel 2019 及 SPSS 26.0 进行数据库的整理与分析。采用前述指标对 DRG 分组效能进行评价, Kruskal-Wallis H 检验水准 $\alpha = 0.05$ 。计数资料以率、百分比表示。

结 果

1. 裁剪前患者住院次均费用分布

结果显示, 各病组变异系数范围为 0.25 ~ 2.87, 有 3 个病组 $CV > 2$, 有 114 个(31%) 病组 $CV > 1$, 说明在病组费用测算过程中, 首先需要处理极值以减少对病组资源消耗水平的估计偏差。各病组患者的次均住院费用呈偏态分布, 除 2 个病组偏度 < 0 外, 各病组的偏度系数均 > 0 , 说明病组普遍呈现正偏态分布。有 17 个病组的偏态系数 > 10 , 右侧长尾明显, 极大值是影响病组费用分布的主要因素。各病组的峰度范围为 $-0.95 \sim 506.12$, 普遍呈现尖峰态, 见图 1。

2. 入组病例住院费用裁剪基本情况

四种裁剪方法对于数据库的裁剪比例各有不同。对数据库裁剪强度最大的是方法 2, 其裁剪比例最高(8.35%), 裁剪去除的极大值与极小值比例也最高, 分别占 5.36% 和 2.99%; 其次为方法 1(5.47%), 其裁剪去除的极大值比例同方法 2, 而裁剪低限较方法 2 宽松, 故裁剪去除的极小值比例低于方法 2, 为 0.11%; 方法 4 上下限各裁剪 2.50%, 裁剪比例低于

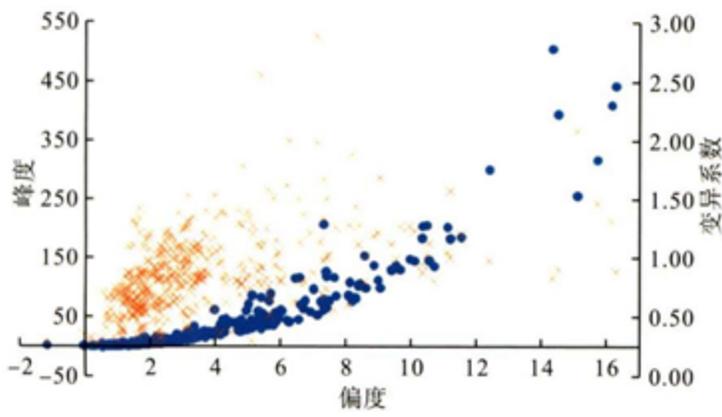


图 1 裁剪前患者住院费用变异系数、峰度与偏度

方法 1 和 2; 裁剪比例最低的为方法 3 (1.30%), 见表 1。

3. 组内异质性评价

结果显示, 入组病组和稳定病组有超过 25% 的病组内病例住院费用差异较大, 方法 1、2 和 3 裁剪后数据库中有 1.63% 的病组内病例住院费用差异较大, 方法 4 裁剪后数据库中仅 0.54% 的病组内 (共计 8 例) 病例住院费用差异较大, 见表 2。

表 1 某市二级医院 2017 - 2019 年历史数据按 CHS-DRG 细分组入组数据次均费用裁剪基本情况

裁剪方法	病例数		裁剪比例 (%)	裁剪掉的极大值病例		裁剪掉的极小值病例	
	裁剪前	裁剪后		例数	百分比 (%)	例数	百分比 (%)
方法 1	526739	497950	5.47	28219	5.36	570	0.11
方法 2	526739	482781	8.35	28219	5.36	15739	2.99
方法 3	526739	519887	1.30	6852	1.30	0	0.00
方法 4	526739	500403	5.00	13168	2.50	13168	2.50

表 2 某市二级医院 2017 - 2019 年历史数据裁剪前后病组内次均费用变异情况

	总病例数	次均费用	病组数	CV > 1 的病组数		CV > 1 的病例数	
				病组数	百分比 (%)	病组数	百分比 (%)
入组病组	528201	3614.93	516	139	26.94	97359	18.43
稳定病组	526739	3603.06	369	114	30.89	97073	18.43
方法 1	497950	3130.48	369	6	1.63	1388	0.28
方法 2	482781	3196.21	369	6	1.63	1388	0.29
方法 3	519887	3427.78	369	6	1.63	1388	0.27
方法 4	500403	3154.25	369	2	0.54	8	0.00

4. 组间异质性评价

经过不同数据裁剪方式, RIV 值从稳定病组的 0.30 最高提升至方法 2 裁剪后的 0.57, 即组间变异能够解释 57% 的总变异。相比之下, 方法 1 对初始值的 RIV 贡献值低于方法 2, 为 0.25; 方法 3 对初始值的贡献仅为 0.08; 方法 4 最差, 为 -0.01。

多个独立样本的 Kruskal-Wallis H 检验结果显示, 稳定病组库和经四种裁剪方式裁剪后病例库各 DRG 组间卡方值所对应的 P 值均为 < 0.001, 说明各组间病例次均费用的差异有统计学意义, 即分组效果较好, 见表 3。

表 3 某市二级医院 2017 - 2019 年历史数据裁剪前后的 DRG 组间差异情况

	DRG 组数	n 个子集住院费用的离均差平方和	总离均差平方和	RIV	χ^2	P
稳定病组	369	7428281798398.53	10680191013101.18	0.30	181954.614	< 0.001
方法 1 裁剪后	369	2076630230555.66	4619449381718.69	0.55	199198.431	< 0.001
方法 2 裁剪后	369	1958348726462.71	4531328477977.04	0.57	207301.721	< 0.001
方法 3 裁剪后	369	4839975774914.43	7839098548823.42	0.38	185287.882	< 0.001
方法 4 裁剪后	369	1520679025206.41	2140425049414.94	0.29	163769.167	< 0.001

5. ROC 曲线评价

用 ROC 曲线评价病例组合分类合理性。结果显示, 三种方法的曲线下面积均大于无效面积 0.5 ($P < 0.05$)。其中方法 1 的曲线下面积最大, 为 0.979, 高于方法 2 的 0.964, 方法 4 的曲线下面积略低, 为 0.695, 详见表 4、图 2。

表 4 三种数据裁剪方法的 ROC 曲线检验

裁剪方法	ROC 曲线下面积	标准误	P	置信区间	
				下限	上限
方法 1	0.979	0.000	< 0.001	0.978	0.979
方法 2	0.964	0.000	< 0.001	0.964	0.965
方法 4	0.695	0.004	< 0.001	0.687	0.703

* : 以方法 3 判断结果为参考变量

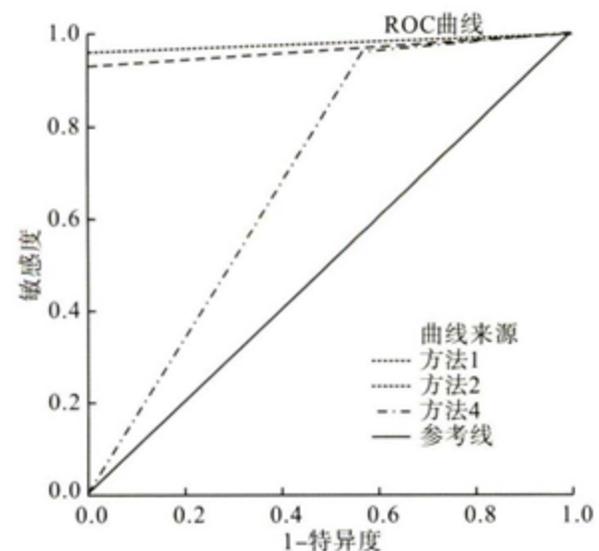


图 2 三种数据裁剪方法的 ROC 曲线

6. 四种方法裁剪结果的综合评价

研究显示,不同的离群值裁剪方法对分组效能的改善程度不同。在有同样的 CV 值改善和 KW 检验通过的情况下,方法 1 和方法 2 在其他分组效能指标上各有所长,方法 1 在尽可能保留原始数据方面优于方法 2,方法 2 对 RIV 值的改善程度略优于方法 1。由于方法 2 对离群值的裁剪没有超过 10%,因此,对于本次分组数据,本研究认为用方法 2 对 DRG 分组效能改善的程度最佳,方法 1 为次优选择,见表 5。

表 5 四种数据裁剪方法结果

裁剪方法	裁剪比例 (%)	CV>1 的病例数 (%)	RIV	ROC 曲线下面积
方法 1	5.47	0.28	0.55	0.979
方法 2	8.35	0.29	0.57	0.964
方法 3	1.30	0.27	0.38	—
方法 4	5.00	0.00	0.29	0.695

7. 数据裁剪对去除高费用病例数病组的影响

数据裁剪对不同病组的影响不同。本研究中方法 1 和方法 2 裁剪的高费用病例数最多(28219 例),分析裁剪去除的高费用病例数占前 10 位的 DRG 组及其占比。结果显示,对“尿路结石、阻塞及尿道狭窄,不伴并发症或合并症”的影响最大,裁剪掉的高费用病例数超过 10%,其次是“脑缺血性疾患”、“呼吸系统感染/炎症”,裁剪掉的高费用病例数占比分别为 6.13% 和 5.81%,见表 6。

表 6 数据裁剪去除高费用病例数占前 10 位的 DRG 组及其占比

序号	DRG 编码	DRG 组名称	超高限病例数	病组病例数	超高限占比 (%)
1	ES29	呼吸系统感染/炎症	2863	49260	5.81
2	DT19	中耳炎及上呼吸道感染	2236	40719	5.49
3	BR29	脑缺血性疾患	1917	31295	6.13
4	IU29	颈腰背疾患	1568	31357	5.00
5	EX29	百日咳及急性支气管炎	1260	22162	5.69
6	GW19	食管炎、胃肠炎	1237	28689	4.31
7	ET29	慢性气道阻塞病	1192	22604	5.27
8	LX15	尿路结石、阻塞及尿道狭窄,不伴并发症或合并症	914	7557	12.09
9	FR49	冠状动脉粥样硬化/血栓/闭塞	666	13309	5.00
10	PU19	足月儿	491	8875	5.53

讨 论

DRG 将几十万,甚至几百万份病例纳入 600 余个病组,故需要评价病例分组的合理性,即分组效能。本研究显示,除个别病组外,绝大部分 DRG 病组的费用数据呈现正偏态分布,数据分布陡峭,因此需去掉离群值再判断分组效能。本研究中费用数据的特征与既往研究一致^[20],采用四种方法裁剪离群值,并通过 CV 值、RIV 值、KW 统计量、ROC 曲线反映裁剪前后 DRG 分组效能。结果显示,四种方法裁剪离群值后对 DRG 分组效能指标均有改善,但改善程度各有不同。

就裁剪强度而言,方法 2 整体裁剪强度最大,其对极大值的裁剪强度与方法 1 相同,对极小值的裁剪强度超过方法 1;其次是方法 4;方法 3 裁剪强度最低,此法只针对 CV>1 的病组裁剪,裁剪比例最小,无论是极大值还是极小值裁剪例数均最少,裁剪后对 RIV 的改善不足,存在极值处理不足的情况。

数据裁剪降低了病组的组内变异, CV 值明显减小,从最初 139 个 CV 值大于 1 的病组(共计 97359 例)最低减少到方法 4 裁剪后的 2 个病组(共计 8 例),起到了 DRG 病组归集的作用。组间变异结果显示,经 KW 检验后,四种方法裁剪后各组间病例次均费用差异均有统计学意义,但经方法 2 裁剪后 RIV 值改善最为明显,较初始值 0.30 已经有了较为明显的提升,方法 1、3、4 的 RIV 值依次降低,且方法 4 对组间变异较初始并无改善作用。

综合 ROC 评价结果显示,以方法 3 作参考变量,方法 1 的曲线下面积最大,方法 2 略低,方法 4 最低,且均大于无效面积。说明由不同方法裁剪后的数据库,各 DRG 组间的住院费用均有统计学差异,进一步说明裁剪后分组结果具有合理性,以方法 1 的裁剪效果最好。

去除离群值病例数前 10 位 DRG 组的分析发现,有 9 个组超高限病例数占比超过 5%,其中包括 1 组占比超 10% 的 LX15 病组,提示医院可能存在住院适应症和诊疗行为不规范问题,提示地方医保局加强关注。

本研究通过对不同数据裁剪方法的比较,显示在特定的数据分布特征下,不同的离群值裁剪方法实现了不同程度的分组效能改善。对离群值进行恰当的裁剪帮助医保合理确定疾病治疗成本,进行精准定价;促使医院规范诊疗行为,加强成本管理,在控制费用的同时保障患者诊疗质量与安全,提高患者就医满意度,实现医、患、保的三方共赢。

离群值处理的效果首先取决于原始数据分布特征,本研究探索不同的数据处理方法后,确定了对于本次数据最适宜的数据裁剪方法,并不意味着该方法适用于所有地区、所有分组方案的 DRG 数据。不同地区和不同类型的医院,应该进一步探索对于不同分布特征的数据的处理方法,以把握 DRG 数据裁剪方法的普遍适用性。

参 考 文 献

- [1] Fetter RB, Shin Y, Freeman JL, et al. Case mix definition by diagnosis-related groups. Med Care, 1980, 18(2 Suppl): 1-53.
- [2] 国家医疗保障局. 国家医疗保障疾病诊断相关分组(CHS-DRG)分组方案. <http://www.nhsa.gov.cn/module/download/downfile.jsp?classid=0&filename=289d4e344c30423b95e33e81ddd086e7.pdf>, 2019-09-30.

- [3] Cots F, Elvira D, Castells X, et al. Relevance of outlier cases in case mix systems and evaluation of trimming methods. *Health Care Manag Sci*, 2003, 6(1) : 27-35.
- [4] 胡佳, 朱晓丽, 郑英, 等. 福建省尤溪县医保支付方式改革及效果研究. *中国卫生政策研究*, 2019, 12(5) : 25-31.
- [5] Reid B, Sutch S. Comparing diagnosis-related group systems to identify design improvements. *Health Policy*, 2008, 87(1) : 82-91.
- [6] 马双, 孟宪芮, 甘犁. 养老保险企业缴费对员工工资、就业的影响分析. *经济学*, 2014, 13(3) : 969-1000.
- [7] 陈吟, 郭默宁, 仇叶龙, 等. 成人对应儿童外科类 DRG 组住院费用影响因素分析. *中国病案*, 2018, 19(4) : 52-57.
- [8] 卢小君. 医疗保险对流动老年人医疗服务利用的影响: 基于倾向得分匹配方法的反事实估计. *中国卫生事业管理*, 2019, 36(9) : 657-660.
- [9] 张宏, 朱立强, 朱东喆, 等. DRGs 在儿童热液烫伤异常费用管理中的应用. *解放军医院管理杂志*, 2014, 21(7) : 644-646.
- [10] 袁丽萍, 高月霞, 许琼琼, 等. 胃癌患者住院费用的 DRGs 病例组合研究. *江苏卫生事业管理*, 2018, 29(12) : 1425-1429.
- [11] Medarevic AP. Describing Serbian Hospital Activity Using Australian Refined Diagnosis Related Groups: A Case Study in Vojvodina Province. *Zdr Varst*, 2019, 13; 59(1) : 18-26.
- [12] 邵慧丽, 宁传英. 基于 E-CHAID 算法胆囊结石患者 DRGs 分组研究. *中国卫生统计*, 2019, 36(4) : 554-555 + 559.
- [13] 简伟研, 卢铭, 胡牧. 北京诊断相关分组(BJ-DRGs) 分组效果的初步评价. *中华医院管理杂志*, 2011, 27(11) : 854-856.
- [14] 朱东喆, 朱立强, 石富胜, 等. 儿童热液烫伤病例 DRGs 分组方法的研究. *中国病案*, 2014, 15(1) : 36-39.
- [15] 付婷辉, 张乐辉, 郭默宁, 等. 对 DRGs 分组方案科学性合理性的分析比较. *中华医院管理杂志*, 2015, 31(11) : 828-830.
- [16] 刘芷辰. DRG 支付体系构建的国际经验借鉴. https://www.sohu.com/a/292058885_133140, 2019-01-29.
- [17] 张晶, 王富珍, 张沛刚, 等. BJ-DRGs 对某女性生殖系统疾病及功能障碍的适用性评价. *中国卫生质量管理*, 2019, 26(4) : 25-27.
- [18] 蒲虎. Kruskal-Wallis 检验原理介绍及其应用. *兴义民族师范学院学报*, 2019(4) : 108-111.
- [19] 姜博文, 马文莉, 李林贵, 等. 宁夏三级医院病例组合分类研究. *卫生经济研究*, 2014(12) : 35-37.
- [20] 蒋柳村, 李佳瑾, 崔欢欢, 等. 基于病案数据的胃癌患者住院费用构成及影响因素分析. *肿瘤预防与治疗*, 2019, 32(11) : 1013-1017.

(责任编辑:刘 壮)

(上接第 350 页)

- [3] Shuang W, Liang X, Jonas JB, et al. Prevalence and Associated Factors of Dyslipidemia in the Adult Chinese Population. *PLOS ONE*, 2011, 6(3) : e17326.
- [4] Kaewprag P, Newton C, Vermillion B, et al. Predictive models for pressure ulcers from intensive care unit electronic health records using Bayesian networks. *BMC Medical Informatics and Decision Making*, 2017, 17(S2) : 65.
- [5] Burnside ES, Rubin DL, Shachter RD. Using a Bayesian network to predict the probability and type of breast cancer represented by microcalcifications on mammography. *Studies in health technology and informatics*, 2004, 107(Pt 1) : 13-17.
- [6] Hughes RE. Using a Bayesian Network to Predict L5/S1 Spinal Compression Force from Posture, Hand Load, Anthropometry, and Disc Injury Status. *Applied Bionics & Biomechanics*, 2017 : 2014961.
- [7] 魏珍, 张雪雷, 饶华祥, 等. 禁忌搜索算法的贝叶斯网络模型在冠心病影响因素分析中的应用. *中华流行病学杂志*, 2016, 37(6) : 895-899.
- [8] Mancini F, Sousa FS, Hummel AD, et al. Classification of Postural Profiles among Mouth-breathing Children by Learning Vector Quantization. *Methods of Information in Medicine*, 2011, 50(4) : 349-357.
- [9] Wang KJ, Makond B, Wang KM. Modeling and predicting the occurrence of brain metastasis from lung cancer by Bayesian network: A case study of Taiwan. *Computers in Biology & Medicine*, 2014, 47 : 147-160.
- [10] Tsamardinos I, Brown LE, Aliferis CF. The max-min hill-climbing Bayesian network structure learning algorithm. *Machine Learning*, 2006, 65(1) : 31-78.
- [11] 张雪雷. 基于禁忌搜索算法的贝叶斯网络在疾病预测与诊断中的应用. 山西医科大学, 2015.
- [12] 中国成人血脂异常防治指南. *中华心血管病杂志*, 2007, 35(5) : 390-419.
- [13] 中国高血压防治指南 2010. *中华心血管病杂志*, 2011, 39(7) : 579-616.
- [14] 董宗然, 周慧. 禁忌搜索算法评述. *软件工程师*, 2010(2) : 96-98.

(责任编辑:邓 妍)