

疾病诊断相关分组 (DRGs) 在肿瘤门诊特殊疾病病种支付中的应用研究

冯海欢* 杨芳 李佳瑾 滕世伟 孙麟

四川大学华西医院医保办公室 四川成都 610041

【摘要】通过多元线性回归模型筛选对医疗费用有统计学意义的影响因素并筛选费用异常数据,利用基于 E-CHAID 算法的决策树模型进行 DRGs 分组,用变异系数、方差减少量及非参数检验验证分组的合理性。通过统计模型剔除 11 条异常数据后共分为 7 个 DRGs 组,经 CV、RIV 及秩和检验验证后证实分组效果合理且较为稳定。相关部门可以门诊特殊疾病为试点逐步推开门诊 DRGs 的应用,分组时需从数据分布特点入手,在合理的数据基础上结合疾病特征、治疗方式等因素进行分组并动态调整,将 DRGs 分组的“事前控制”转变为“事前测算—事中控制—事后调整”,提高分组的综合性和实用性。

【关键词】疾病诊断相关分组;病种付费;肿瘤系统

中图分类号:R197 文献标识码:A doi:10.3969/j.issn.1674-2982.2018.05.011

Study on the application of DRGs in the payment of special disease outpatients in tumor system

FENG Hai-huan, YANG Fang, LI Jia-jin, TENG Shi-wei, SUN Lin

West China Hospital of Sichuan University, Chengdu Sichuan 610041, China

【Abstract】The multiple linear regression models were used to screen the statistically significant risk factors of medical expenses and abnormal data. DRGs grouping was performed using the decision tree model based on the E-CHAID algorithm with the variation, variance reduction and non-parametric test coefficients to verify the rationality of groups for preferred results. After eliminating 11 abnormal data by the statistical model, results found that DRGs can be divided into 7 groups. After verifying CV, RIV and rank-sum test, the results showed reasonable and stable DRGs grouping. The relevant departments can gradually carry out the application of DRGs in outpatient payment reform by setting special diseases as an experimental data. When grouping, it is necessary to start with the characteristics of data distribution, combining and dynamically adjusting DRGs according to factors such as disease characteristics and treatment methods based on reasonable data. The DRGs comprehensiveness and practicability can be improved by changing “pre-control” to “pre-estimate, process-control, and post-adjustment”.

【Key words】DRGs; DRGs-based payment system; Tumor

为缓解“看病难、看病贵”问题,某市(以下称“C市”)医保将 39 种病情相对稳定、需长期在门诊治疗的疾病纳入门诊特殊疾病(以下简称“门特”),患者与医院进行按项目结算,医保经办机构对医院则按照病种大类进行定额拨付,结算周期内患者实际费用达到或超过各病种支付标准的,经审核后按病种支付标准清算拨付,超支部分由医院按照一定比例

进行分担。这种单一的病种付费制度虽然是国际上较为认可的付费机制^[1],但覆盖面有限且付费标准过于粗犷,如肿瘤患者不管肿瘤类型、有无转移及治疗方式如何,均采用统一的付费标准。因此难以保证医院给患者提供科学有效、有针对性的医疗服务,导致医疗质量与医疗消耗的分离。

DRGs(Diagnosis Related Groups)是一种以医疗

* 基金项目:成都市医保基金监控评价统计分析模型建立基金(JH2014093)

作者简介:冯海欢,女,统计师,硕士,主要研究方向为生物统计与医保政策。E-mail:53841484@qq.com

通讯作者:孙麟。E-mail:331221241@qq.com

资源消耗强度为分组轴心,应用统计控制理论的原理将病人归类的新方法,其基本思想是以病例诊断和(或)操作等特征相似的诊断相关疾病进行归类组合,通过对每个组设定支付标准的预付制度,以控制医疗费用过快增长。DRGs 经超大样本量研究制定,综合考虑病人病情、疾病预后、治疗难度及医疗资源消耗等各个方面,将病例进行分类组合形成若干个诊断相关组,使得每一组卫生资源消耗都相近,同质性较高^[2],可以有效平衡医疗质量与费用。因此,DRGs 不仅能为医保支付标准的确定提供科学依据,而且还可以为医院实现有序和有效的质量管理和优质服务提供可以借鉴的宝贵经验和指导方法。

1 对象与方法

1.1 对象

本研究以 C 市某综合性三甲医院(以下简称 A 医院)2016 年 1—6 月医保结算数据中的肿瘤门特数据为例,探讨 DRGs 思想在按病种付费制度中的应用,为按病种支付方式为主的医保支付方式改革提供有理可循的方法与思路。本次研究纳入分析的共有 1 198 名肿瘤患者的门特记录,每条记录中包括患者个人信息、门特申请信息、原始费用明细及医保支付情况。

1.2 方法

DRGs 是基于医疗资源的消耗,本文选择医疗费用代表医疗资源消耗。利用个人信息匹配病案首页中该病人的疾病情况,如肿瘤类型、是否手术、是否

有转移。运用 SAS 9.4 对数据进行逻辑差错和预处理及数据建模。首先利用多元线性回归模型筛选对医疗费用有统计学意义的影响因素并建立费用预测模型,通过预测模型筛选费用异常数据,从而标定相对合理的数据;最后利用基于穷尽卡方自动互动检验法(Exhaustive Chi-square Automatic Interaction Detector, E-CHAID)的决策树模型进行 DRGs 分组并对各组的合理费用进行模拟预测。针对分组效果则根据数据分布情况采用合理的指标检验组间异质性及组内同质性。

2 结果

2.1 基本情况

如表 1 所示,A 医院不同特征的肿瘤患者的医疗费用仅在不同的治疗方式下出现有统计学意义的差异,其他特征的费用差异均无统计学意义。从该表可以看出,不同特征肿瘤门特患者半年费用的均数基本达到一万元以上,即年均费用至少为 2 万元,而 C 市医保对于 A 医院肿瘤门特患者的年均支付费用仅为 13 000 元,按照 90% 的政策补偿比例测算,肿瘤门特患者一年的医疗费用需控制在 14 500 元以下,医院才不会面临分担风险。A 医院在保证医疗质量与安全的前提下,每年都面临肿瘤超额而分担的扣款,不仅影响到医院的正常运营而且使医保基金得不到有效使用。因此针对现有门特定额付费制度,以 DRGs 为途径对病种费用进行精细化管理显得尤为必要。

表 1 A 医院不同特征肿瘤门特患者变量赋值及费用比较

特征变量	因素	例数	构成比(%)	总费用(元)		检验统计量	P	
				M	SD			
个人信息	性别	1 = 男	600	50.08	12 273	12 860	1.149	0.284
		2 = 女	598	49.92	11 644	12 987		
	年龄	原始值纳入	1 198	/	11 959	12 923	/	/
疾病情况	是否转移	0 = 否	1 113	92.9	11 849	12 989	0.715	0.398
		1 = 是	85	7.1	13 398	11 992		
	是否手术	0 = 否	186	15.53	11 546	13 296	0.007	0.934
		1 = 是	1 012	84.47	12 035	12 858		
治疗方式	0 = 维持性治疗	264	22.00	1 827	1 958	107.921	<0.001	
	1 = 化疗	763	63.70	13 894	11 865			
	2 = 放疗	36	3.00	9 284	13 873			
	3 = 放化疗	135	11.30	21 553	17 751			

2.2 影响因素筛选与合理数据标定

本文以个人半年总费用为因变量,以性别、年

龄、是否转移、是否手术、药品费用构成、治疗方式为自变量(变量纳入形式见表 1)建立多元线性回归模

型,采用逐步回归法筛选自变量。其中半年总费用成偏态分布,故转换为对数值纳入模型。药品费按照《国家基本医疗保险、工伤保险和生育保险药品目录》进行分类。由于本研究数据限制,未能包含患者合并症并发症变量,但分类后的药品费可间接反映患者的合并症及并发症。由于患者发生费用的累计月份数有差别导致累计费用有较大差异,因此本文以各类费用占比纳入分析。

从表 2 可以看出,模型拟合效果较好 ($R_{adj}^2 = 0.976$)。治疗方式和年龄对肿瘤患者医疗费用的影响有统计学意义,以治疗方式的影响效应最大,其余变量对医疗费用的影响不具有统计学意义。以该模型为预测模型,95% 置信区间为临界值筛选异常数据,共筛选 11 条实际费用高出预测临界值的异常记录,其余记录则标定合理数据进一步进行 DRGs 分组。

表 2 各变量对肿瘤患者半年总费用的影响效应

变量	偏回归系数	标准误	标准偏回归系数	<i>t</i>	<i>P</i>
常数	3.499	0.393		8.902	<0.001
治疗方式	2.191	0.136	0.843	16.167	<0.001
年龄	0.051	0.006	0.467	8.961	<0.001

注: $R^2 = 0.981, R_{adj}^2 = 0.976, F = 181.246, P < 0.001$

2.3 DRGs 分组及模型评价

本文采用基于 E-CHAID 算法的决策树模型进行

DRGs 分组。E-CHAID 算法是 CHAID 的改进算法,广泛适用于定类、定序和定量数据的分析,是一种以目标最优为依据,具有目标选择、变量筛选和聚类功能的分析方法。它是根据经过筛选的解释变量对数据中给定的因变量进行最优分割,通过卡方检验/方差分析是否具有显著性进行自动判断分组,使得亚组之间的差异(不纯度)最大而亚组内差异尽可能的小(纯度尽可能大)。E-CHAID 在 CHAID 的基础上在选择最佳分组变量时采用“将分组进行到底”的策略,比 CART 算法和 CHAID 算法更适合于建立 DRGs 模型。^[3]模型以个人总费用对数值为因变量,以回归模型中对费用有影响意义的年龄、治疗方式为自变量,同时考虑到不同部位的肿瘤具有不同的特点,因此将肿瘤所属部位、年龄、治疗方式为自变量进行分类。

2.3.1 DRGs 分组情况

经过对合理数据进行标定后,肿瘤门特的医疗费用共分为 7 个 DRGs 组,分组变量中治疗方式、年龄和肿瘤系统具有统计学意义。7 个组中例数最多的为第 5 组,即年龄大于 47 岁进行单纯化疗或放疗的患者,占总例数的 62%。费用最低的为甲状腺系统或泌尿系统的维持性治疗,人均半年总费用不到一千元;费用最高的为消化系统或呼吸系统或泌尿系统或唇、口腔、咽系统同时进行放化疗治疗,人均半年总费用高于两万元。

表 3 DRGs 分组结果及模型评价指标

DRGs	组合	例数	集中趋势		离散趋势		变异系数		<i>RIV</i>	秩和检验	<i>P</i>
			<i>m</i>	<i>M</i>	<i>SD</i>	四分位间距	四分位间距/ <i>M</i>	标准差/ <i>m</i>			
1	维持性治疗,甲状腺或泌尿系统	65	785	481	775	4075	8.480	0.987			
2	维持性治疗,消化或乳腺或淋巴造血系统	123	2 163	1 788	2 246	12 739	7.124	1.038			
3	维持性治疗,其他系统	76	2 177	1 805	1 866	8 343	4.623	0.857			
4	化疗或放疗,年龄 ≤47	53	11 073	7 432	11 615	47 574	6.402	1.049	23.18%	501.326	<0.001
5	化疗或放疗,年龄 >47	746	13 872	10 180	12 004	80 290	7.887	0.865			
6	同时放化疗,消化或呼吸或泌尿或唇口腔和咽系统	72	26 612	22 176	18 796	81 773	3.687	0.706			
7	同时放化疗,其他系统	63	15 773	11 589	14 588	78 155	6.744	0.925			
	分组前	1 198	11 960	8 022	12 923	83 539	10.414	1.081			

2.3.2 模型评价

(1) 组内同质性

DRGs 病例组合的原理即尽量增大组合内的“同质性”。一般通过计算变异系数(Coefficient of Variation, CV)对每个 DRGs 组的组内差异进行评价。CV

是一个数据集的变异指标与其平均指标的比值,越小表示组内变异越小,同质性越好。变异指标可以为全距、平均差或标准差等^[4],传统计算公式为 $CV = \text{标准差} / \text{均数}$ 。由于费用数据呈偏态分布,本文首次尝试用描述偏态分布的四分位间距代替标准差,

用中位数代替均数来衡量组内变异程度。

表3所示,本文所使用的两种变异系数计算方法均显示出,分组后各DRGs组的变异程度均小于分组前数据集的变异程度,说明分组后组内同质性较分组前有所提高。

(2) 组间异质性

一般采用方差减少量(Reduction in Variance, RIV)计算一个数据集被分解为 $n(n \geq 2)$ 个子集后子集间的变异占总变异的比值来衡量各检测集DRGs组间的异质性。RIV值越大则组间异质性越强,分组结果越好。本研究分组后RIV值为23.81%,可以认为本次分组效果较好。

(3) 分组模型成立

由于医疗费用成偏态分布,本研究采用Kruskal-Wallis H检验DRGs分组后模型是否成立。结果显示各组间的费用不全相同,且差异具有统计学意义($H = 501.326, P < 0.001$),可以认为分组模型成立。

2.4 付费标准参考值制定

医保经办机构可以各DRGs组的集中趋势指标与离散趋势指标为参考进行付费标准的范围界定。如首先以各组合的费用中位数为参考,结合当地医保患者的政策补偿比例和基金收支情况制定病种付费标准(表4)。再以总费用的75%加1.5倍四分位间距或者几何均数加减2倍或3倍标准差作为限外病人的控制限^[5-7],超过控制限的患者标定为异常数据并进一步分析,不仅可以消除医疗费用离散度过大的异常值影响,还可以通过了解限外患者的疾病严重程度及医疗资源消耗情况,为病种付费标准的动态调整提供科学依据。相比于分组前,分组后各组的均数或中位数都呈现出层次性,能够体现出不同DRGs组别的临床治疗及疾病情况特征,不仅支付标准更具有聚焦性,而且更容易识别限外病人等异常数据,有利于医保经办机构进行有针对性的管理。

表4 DRGs组费用偿付标准参考(元)

组合	例数	均数	中位数	费用上限
维持性治疗,甲状腺或泌尿系统	65	706	432	6 031
维持性治疗,消化或乳腺或淋巴造血系统	123	1 947	1 609	18 658
维持性治疗,其他系统	76	1 959	1 624	12 733
化疗或放疗,年龄 ≤ 47	53	9 966	6 688	71 700
化疗或放疗,年龄 > 47	746	12 485	9 162	117 755
同时放化疗,消化或呼吸或泌尿或唇口腔和咽系统	72	23 951	19 958	128 357
同时放化疗,其他系统	63	14 196	10 430	116 157
分组前	1 198	10 764	7 220	120 850

注:按照90%政策补偿比例计算

3 结论与建议

3.1 结论

医保支付标准的确定需依据“大数法则”,传统的病种付费标准仅仅是粗略地以一定时间段内某病的平均值为参考,并没有考虑到数据的实际分布情况及疾病特征、病人自身特点等因素,因此对于医院和医保经办机构都难以做到精细化管理。DRGs作为一种理想的付费方式,其研究领域大多集中在住院病例^[8],近年来某些特殊疾病及专项重大疾病如脑梗、白血病、心脏病等疾病也是DRGs的研究热门^[9,10],但鲜有针对门诊病例的研究。

本文首次将需要长期维持治疗的肿瘤门特患者纳入研究,采用与成本密切相关的医疗费用测算各组合的支付标准,从数据本身的分布特点入手,利用

统计学方法进行费用的影响因素筛选,通过科学的方法挖掘异常数据,在合理的数据基础上结合疾病特征、治疗方式等因素对费用进行分组,不仅扩大了DRGs思想在医保支付方式改革的应用领域,而且非参数Kruskal-Wallis H检验、RIV值测算均显示组间异质性较强,CV测算结果显示组内医疗费用消耗的同质性较好,说明了病例组合方案的可靠性和合理性。此外,本文创新性纳入了反映疾病治疗过程的变量(治疗类型)并显示出对医疗费用有较强的影响,不仅能够弥补门诊患者由于诊断不完善导致的疾病严重程度不明确的缺点,而且将DRGs分组的“事前控制”转变为“事前—事中控制”,使分组的综合性和有效性更强。因此本文介绍的方法可推广于门诊的其它系统疾病,最终形成一套包括所有疾病在内的、开放性的和动态的病种费用标准体系。

3.2 建议

3.2.1 以门诊特殊疾病为试点逐步推广门诊 DRGs 的应用

各地由于医保政策不同,门诊支付方式有所差异。C 市门诊仅将部分需要长期治疗的慢性非传染性疾病纳入医保支付,该部分疾病的患者人群和疾病状态较为稳定,与普通门诊患者相比更适合于 DRGs 付费。全国有部分地区实现了门诊统筹,支付方式多为按项目付费或包干定额付费,这种支付方式不仅面临医保基金浪费的风险,而且没有考虑到病人的实际就医需求。由于门诊的特殊性,其数据比住院数据更为庞大和复杂,尤其是疾病诊断规范性往往得不到保证,病人的流动性强,病历信息有限。门特数据则一定程度上规避了这类问题。以 C 市为例,所有纳入门特的疾病均为大类病种定额付费,如糖尿病,甚至都没有区分 I 型还是 II 型,高血压也没有区分级别。当病种发生超定额标准时仅体现出该大类病种的就医需求与医保基金耗用不平衡,医院难以确定超标准的详细原因,医保经办机构面临医保基金安全可持续性管理时也没有科学可靠的抓手。因此以门特患者为试点逐步推广门诊其他相对稳定的病种进行 DRGs 付费,对大类病种进行精细化管理具有一定的现实意义和可操作性。

3.2.2 结合当地疾病谱和不同级别医院功能定位进行动态调整

病种医疗费用标准的制订是 DRGs 分组结果的一个重要应用,但也是一个艰巨的任务。医疗费用本身受多种因素影响,除病人病情及病人的个体特征外,还有很多客观因素如当地物价指数,医院规模及所在区域社会经济特征等。这些因素相互交叉影响,难以分开,导致不能用统一的统计模型进行描述,因此,在制定标准费用的过程中需要考虑到临床诊断的同时,还要考虑到病人需求、治疗过程及当地疾病发展趋势,并结合医改政策对医院的影响前瞻性预测医院病种收治结构的变化,将 DRGs 付费机制完善为“事前测算,事中控制,事后调整”,一方面能够兼顾医疗费用与医疗质量的平衡,另一方面能够提高医保基金的风险防范能力并提高基金的有效使用率。

3.3 需进一步研究的内容

病例组合是基于大样本数据,采用统计学方法拟合的决策树模型对医疗资源消耗的病例进行分类的过程。如果样本量太小,则统计结果没有代表意义,结果不稳定。本研究虽然采用了 1 198 份有效样本,但仅

为同一家医院的半年数据,不同类型的肿瘤患者代表性有限,下一步可通过纳入其他医院及更长时间段内的肿瘤门特患者数据来扩大样本含量和样本信息,提高研究的可信度和效度。另外,用医疗成本进行 DRGs 分组能够代表疾病的真实消耗,但现阶段我国根据成本确定 DRGs 支付标准仍有相当大的难度^[11],可进一步完善基于 DRGs 系统的成本核算体系、成本信息数据库^[12],对相应的支付标准进行调整和修正,使 DRGs 在病种支付方式改革的应用不断得到完善。

作者声明本文无实际或潜在的利益冲突。

参 考 文 献

- [1] 雷海潮. 中国卫生总费用的未来趋势及相关政策思考[J]. 中国卫生事业管理, 1999(3): 137-140.
- [2] Issel L M, Anderson R A. Avoidable costs of comprehensive case management[J]. Health Care Management Research, 2001, 24(3): 64-72.
- [3] 杜剑亮, 刘骏峰, 陈倩. 不同决策树算法建立 DRGs 模型的差异[J]. 中国病案, 2014, 15(7): 38-41.
- [4] 杜剑亮, 刘骏峰, 陈倩. 常用 DRGs 模型的评价方法[J]. 中国病案, 2014, 15(8): 36-37.
- [5] 高子厚, 万崇华, 蔡乐, 等. 按 DRGs 组合方式制定消化系统疾病患者住院费用的研究[J]. 中国卫生统计, 2006, 23(4): 323-325.
- [6] 徐勇勇, 张音, 潘峰, 等. 基于我国病案首页的病例组合方案与病例组合指数[J]. 中华医院管理杂志, 2001, 17(1): 34-36.
- [7] 黄慧英. 诊断相关分类法在北京地区医院管理中的可行性研究[J]. 中华医院管理志, 1999, 10(3): 131-149.
- [8] 简伟研, 胡牧, 张修梅. 诊断相关组(DRGs)的发展和应[J]. 中华医院管理杂志, 2011, 27(11): 817-820.
- [9] Cleary Ml et al. The future of case mix in emergency medicine and ambulatory care[J]. Med, 1994 sep 5; 161(suppl): 530-533.
- [10] Ashby R, Jelinek G. Case mix classification for outpatient services based on episodes of care[J]. Med J, Aust 1993 aug 2; 159(3): 213.
- [11] 林晨蕾. 美国 DRGs 支付制度对我国医疗保险支付方式的启示[J]. 中外企业家, 2010(10): 115-116.
- [12] 曾雁冰, 林鹏, 方亚. 基于 CHAID 算法的病毒性肝炎患者 DRGs 分组研究[J]. 中国卫生统计, 2015, 32(3): 514-517.

[收稿日期:2017-05-10 修回日期:2017-12-24]

(编辑 薛云)