

# 四川省三级医院慢性阻塞性肺疾病患者参加医疗保险类型与其住院时间的关系

## ——基于分位数回归模型的实证分析

蔡瑞烈\* 潘杰

四川大学华西公共卫生学院 四川成都 610041

**【摘要】**目的:分析慢性阻塞性肺疾病(COPD)患者参加医疗保险类型与其住院时间的关系,对患者平均住院日的降低和医疗保险支付方式的完善提供参考。方法:利用2016年第四季度在四川省三级医院就诊的36 329例COPD患者病案首页信息,采用分位数回归模型,分析患者不同参保方式与其住院天数的关系。结果:不同参保方式患者的住院时间存在统计学差异,且这种差异大部分是由参保方式不同造成的。全公费患者住院天数最长,新型农村合作医疗参保患者最短,这种差异随住院时间分位数的提高而变大。结论:不同参保方式的COPD患者住院时间存在差异。全公费和城镇职工医疗保险参保患者的住院天数相对更长的可能原因是医院过度提供医疗服务;新型农村合作医疗参保患者的住院时间最短,可能原因是医院将其住院床位转移给了全公费和城镇职工医疗保险患者;不同医疗保险主办方监管力度不同也可能是造成差异的原因。

**【关键词】**慢性阻塞性肺疾病;住院时间;医疗保险;分位数回归

中图分类号:R197 文献标识码:A doi:10.3969/j.issn.1674-2982.2018.11.004

### The relationship between the types of medical insurance and the length of stay of COPD patients in tertiary hospitals of Sichuan province: An empirical analysis based on quantile regression model

CAI Rui-lie, PAN Jie

West China School of Public Health, Sichuan University, Chengdu Sichuan 610041, China

**【Abstract】** Objective: The goal of this study is to analyze the relationship between the types of medical insurance chosen by patients with chronic obstructive pulmonary disease (COPD) and their length of stay in hospitals (LOS), and thus to provide policy advice for improving medical insurance payment methods and reducing the average length of stay in hospitals (ALOS). Methods: Information on the first page of medical records of 36,329 patients with COPD in tertiary hospitals of Sichuan Province in the fourth quarter of 2016 are collected, and the quantile regression model is used to analyze the correlation between different types of medical insurance and length of stay. Results: There is a statistically significant difference in the length of stay between patients with different types of medical insurance, and most of this difference is caused by the difference in insurance types. Patients with free medical care have the longest LOS while patients with the new rural cooperative medical scheme (NRCMS) have the lowest, and the difference increases with the quantile. Conclusions: There are differences in the length of stay between different types of medical insurance. The reason for the relatively longer LOS for patients with free medical care and urban employee basic medical scheme (UEBMS) is that hospitals provide excessive medical services. On the other hand, the NRCMS patients have the shortest LOS possible because hospitals may transfer their inpatient beds to free medical care and UEBMS patients. At the same time, the supervision of different medical insurance sponsors may also be the cause of the difference.

**【Key words】** Chronic obstructive pulmonary disease; Length of stay; Medical insurance; Quantile regression

\* 作者简介:蔡瑞烈,男(1993年—),硕士研究生,主要研究方向为卫生统计方法在卫生政策的应用。E-mail: 1060003418@qq.com  
通讯作者:潘杰。E-mail: panjie.jay@hotmail.com

## 1 引言

慢性阻塞性肺疾病 (COPD) 是一种以气流阻塞为特征的常见慢性疾病,主要表现为慢性支气管炎和(或)肺气肿<sup>[1]</sup>,严重影响患者的劳动力和生活质量。COPD 患者人数众多,2000—2014 年我国 40 岁以上成人 COPD 患病率为 8.4% ~ 10.1%。<sup>[2]</sup>在我国所有疾病中,COPD 的疾病经济负担位列第四。<sup>[3]</sup>在如此高患病率和高负担的情况下,COPD 患者对医疗资源的利用效率显得尤为重要。

平均住院日作为三级医院评审的核心指标之一<sup>[4]</sup>,是衡量一家医院医疗质量和管理水平的重要指标,综合反映住院患者使用医院医疗服务的水平。过长的住院时间不仅增加患者的经济和精神负担,还会浪费宝贵的医疗资源。<sup>[5]</sup>在床位紧缺的三级综合医院,平均住院日通常难以受患者主观因素决定,因此其一定程度上反映了医院是否存在过度提供医疗服务的情况。

基于以上背景,本文的研究问题是:不同参保方式的 COPD 患者医疗服务资源的利用有无差别,医院是否对一些参保类型的患者存在过度提供医疗服务的情况?目前国内外大量学者已对住院时间或者医保参保类型进行研究。如研究不同疾病患者或不同特征患者住院时间的影响因素。<sup>[6]</sup>部分研究在讨论住院时间时,没有考虑住院时间数据呈现正偏态分布的特点,盲目使用正态分布假设下统计描述和统计推断方法,做出不恰当推断。因此,有学者建议在研究中使用中位数描述住院日的平均水平。<sup>[7]</sup>研究住院时间和参保方式的文献多采用构成比分析和单因素分析,不能有效控制其他混杂因素的影响,如疾病因素。多因素分析的研究使用的方法多为最小二乘法回归或 logistic 回归,只能估计相关因素的平均效应。对于研究结论,多数研究认为医保患者比未参保患者的住院时间更长,但也有研究发现参保降低平均住院日。<sup>[8]</sup>目前,医疗保险对平均住院日的影响尚存争议。另一方面,国外文献相比国内文献研究更为深入,主要定量分析某种治疗方法<sup>[9-10]</sup>或某一特定疾病<sup>[11]</sup>住院时间的各种影响因素,以及不同医院类型住院时间差异<sup>[12]</sup>,使用方法也不局限于描述性分析和单因素分析,使用了包括 Poisson 回归、广义线性回归、分位数回归等更多类型的定量分析方法,但相关研究对象往往仅限于某一家医院患者,难以避免具体医院的选择性偏倚。目前,

尚未见对住院时间与医疗保险类型关系展开系统研究的文献。

本研究使用四川省全省住院病人信息,运用分位数回归方法,研究 COPD 患者参保方式与住院时间的关系,为医院提高医疗质量和管理水平,以及医疗保险支付方式的完善提供参考。

## 2 资料和方法

### 2.1 资料来源

数据来源于四川省西医医院病案首页数据,时间为 2016 年第四季度。研究对象为三级医院中入院诊断为 COPD 的患者 (ICD-10 编码为 J44),共计 40 984 例,剔除住院时间为 0、参保方式缺失、转院和再住院、性别缺失观测的病例,最终得到总样本 36 329 例,提取患者的住院信息,包括病人的人口学特征和住院情况,如性别、年龄、现住址、参保方式、出入院情况和并发症等,此外,还结合统计年鉴数据,收集患者所在医院的相关特征。

本研究的因变量是住院时间 (LOS),反映患者医疗服务利用时间的长短,在纳入观测时,剔除了住院时间为 0 的观测值 (3 例)。对于本研究关注的自变量,即参保方式,共计包括 8 种,鉴于样本量的因素,将医疗救助、商业医疗保险和其他社会保险合并为其他医疗保险,同样,剔除了参保方式缺失的观测值 (2 112 例)。综合上述的研究和机构的信息<sup>[13-16]</sup>,本研究最终纳入的控制变量有性别、年龄、职业、疾病严重程度、离院方式、手术、输血、感染、医院等次、医院床位数、医院类别、医院经济类型 (表 1)。

表 1 各变量定义及统计描述表

变量	变量定义	占比/%
住院时间 均值 (SD)/中位数	住院的时间/天	11.86(9.74)/10
参保方式	1 = 新型农村合作医疗	31.73
	2 = 城镇职工基本医疗保险	36.88
	3 = 城镇居民基本医疗保险	20.58
	4 = 全公费	1.15
	5 = 全自费	5.26
	6 = 其他医疗保险	4.40
女性	二分类,1 = 女性	33.68
年龄 均值 (SD)	年龄 (岁)	72.84(10.46)
疾病严重程度	二分类,1 = 严重	45.16
离院方式	1 = 医嘱离院	87.09
	2 = 非医嘱离院	9.94
	3 = 死亡	1.79
	4 = 其他	1.18

(续)

变量	变量定义	占比/%
是否手术	二分类, 1 = 是	13.01
是否输血	二分类, 1 = 是	2.21
是否发生感染	二分类, 1 = 是	15.61
医院等次	二分类 0 = 甲等 1 = 乙等	44.00 56.00
床位数 均值(SD)	实际床位数(床)	1 091.38(675.68)
医院类别	1 = 综合医院 2 = 精神病医院 3 = 传染病医院 4 = 职业病医院 5 = 妇幼保健院	90.00 6.00 2.00 1.00 1.00
医院经济类型	1 = 公立医院 2 = 非公立医院	99.00 1.00
职业	哑变量, 共 12 个	—
位置	哑变量, 共 28 个	—
观测数/人	—	36 329
机构数/个	—	100

注:SD(Standard Deviations) 指标方差

四川省不同市/县的经济发展情况、医疗卫生服务能力、医疗卫生质量是存在差异的,这些因素是会影响患者住院时间的,为了解释这些因素的原因,加入了按地级市及地级市辖县划分的地区哑变量,共 28 个,以识别可能存在的固定效应。

还有其他研究提示教育程度和收入也可能影响患者住院时间<sup>[17]</sup>,理论上,也应该纳入这些控制变量,但是这些信息无法获取。此外,转院的患者和再住院患者的住院时间应该比数据显示的更长,但是由于无法获取转院之后的信息以及无法识别再住院后的观测,所以这部分患者也被剔除(2 539 例)。

## 2.2 分析方法

### 2.2.1 分位数回归

传统的线性回归模型通常估计自变量和整个范围的因变量之间的关系,即自变量 X 对于因变量 Y 均值的影响。其中最基本的方法是使用普通最小二乘法(OLS)来估计回归系数,如果满足高斯—马尔科夫假定,那么回归系数的最小二乘估计为最优线性无偏估计(BLUE),如果进一步误差服从正态分布,那么回归系数为最小方差无偏估计(MVUE)<sup>[18]</sup>但是实际中,这些假设往往很难完全得到满足,此时最小二乘将不会得到上述优良性质,且稳健性会很差。为了弥补普通最小二乘(OLS)在回归分析中的缺陷,Koenker 和 Bassett (1978)提出了分位数回归的思想。<sup>[19]</sup>

对于随机变量 Y 的一个随机样本  $\{y_1, y_2, y_3, \dots, y_n\}$ , 它的中位数线性回归一般采用最小一乘(least absolute, LA)准则,或最小绝对偏差和(minimum absolute deviations)<sup>[20]</sup>,即:

$$\min \xi \sum |y_i - \xi| \quad (1)$$

记 Y 的第  $\tau$  分位数为  $Q(\tau) = \inf\{y: F(y) \geq \tau\}$ , 中位数回归其实是分位数回归的特例( $\tau = 1/2$ ),而  $\tau$  分位数的样本分位数线性回归则是要求满足:

$$\min_{\beta(\tau) \in R^k} \sum \rho_\tau(y_i - x_i' \beta(\tau)) \quad (2)$$

其中对于任意  $0 < \tau < 1$ , 检验函数  $\rho_\tau(\mu)$  为:

$$\begin{aligned} \rho_\tau(\mu) &= (\tau - I_{(\mu < 0)})\mu \\ &= \begin{cases} \tau\mu & \mu \geq 0 \\ (\tau - 1)\mu & \mu < 0 \end{cases} \end{aligned} \quad (3)$$

所以(2)式的展开式为:

$$\begin{aligned} \min_{\beta(\tau) \in R^k} \left[ \sum_{(i: y_i \geq x_i' \beta(\tau))} \tau |y_i - x_i' \beta(\tau)| \right. \\ \left. + \sum_{(i: y_i < x_i' \beta(\tau))} (1 - \tau) |y_i - x_i' \beta(\tau)| \right] \\ \tau \in (0, 1) \end{aligned} \quad (4)$$

在线性条件下,给定 x 后, Y 的  $\tau$  分位数函数为:

$$Q_y(\tau | x) = x' \beta(\tau) \quad \tau \in (0, 1) \quad (5)$$

对于不同的分位数  $\tau$ , 都能得到一个分位数函数,因此随着  $\tau$  从 0 到 1, 就能得到 y 在 x 上的条件分布的轨迹,即一簇曲线,而不像一般线性回归只得到一条曲线。<sup>[21]</sup>

由此可见,百分位数回归其实是一种加权最小一乘(weighted least absolute, WLA),其优点表现为以下几个方面:(1)它对模型误差项不做任何假定, $\mu$  不需要服从正态分布,整个回归模型就具有很强的稳健性,这也正是它相较 OLS 的优势。(2)分位数回归没有使用一个连接函数来描述因变量均值和方差的相互关系,因此它具有较好的弹性性质。(3)分位数回归是对不同  $\tau$  值下一系列的回归,对数据中的异常点具有良好的抗性。(4)不同于 OLS,分位数回归对因变量具有单调变换性。(5)分位数估计出来的参数具有在大样本理论下的渐进优良性。<sup>[19]</sup>

因此,根据纳入的变量,本研究构建的分位数回归方程为:

$$\begin{aligned} Q_{LOS}(\tau) &= \alpha(\tau) + \beta(\tau) controls \\ &+ \gamma(\tau) insurance + \varepsilon(\tau) \end{aligned} \quad (6)$$

其中,LOS 代表住院时间, $\alpha$  代表截距,insurance

代表参保方式, controls 代表控制变量,包括性别、年龄、职业、疾病严重程度、离院方式、是否手术、是否输血、是否感染、医院等次、医院床位数、医院类别、医院经济类型、地区,  $\tau$  代表分位数,  $\varepsilon$  代表误差项。

### 2.2.2 分位数分解

本研究采用 Machado 和 Mata(2005)所使用的分解工资分布变化的方法(MM2005)。<sup>[22]</sup>这种方法是在条件分位数回归的基础上,借助概率积分转换得到边际密度函数的一致估计,再通过随机替换构造反事实分布,以解释产生差异的成因。

利用 MM2005 分位数分解方法的关键是进行反事实分析(the counter-factual analysis),本研究关心的反事实分析即是,如果某一种参保方式(如城镇职工医保)按照其他另一种参保方式(如新农合)的分位数回归参数决定住院时间的话,这一种参保方式(城镇职工医保)的分布会如何?这里定义反事实分布为  $F(y^* | Z^{UE}, \beta^N)$ ,其中  $Z^{UE}$  表示影响城镇职工医保患者住院时间的变量分布,  $\beta^N$  表示影响新农合患者住院时间的变量在每个分位数上的回归参数,  $y^*$  表示如果城镇职工医保患者按照新农合患者的分位数回归参数决定住院时间的话,城镇职工医保患者反事实住院时间的长短。其具体计算步骤为:(1)从均匀分布  $U[0,1]$  中随机抽取一个样本容量为  $l$  的样本  $\{u_1, u_2, u_3, \dots, u_l\}$ 。(2)在新农合患者样本中,分别以  $\tau = u_1, u_2, u_3, \dots, u_l$  做分位数回归。(3)从城镇职工医保患者样本中有放回的重复随机抽样,抽取一个子样本,表示为  $Z^{UE}$ 。(4)把(2)中得到的分位数回归参数和(3)中得到的城职保患者子样本变量分布相结合,得到一个新的样本,即反事实分布样本  $y_\tau^* = Z^{UE} \beta_\tau^N$ 。

在反事实分析基础上进行分位数分解,其方法是将每个分位数上的城乡收入差异分解为两个部分,一部分是由参保方式不同引起的(即分位数回归参数不同引起的),另一部分是由两个样本特征变量分布不同引起的(即影响因素变量值的不同引起的)。假定  $\tau$  分位数下城镇职工医保患者住院时间、反事实住院时间和新农合患者住院时间分别为  $y_\tau^{UE}$ 、 $y_\tau^*$ 、 $y_\tau^N$ 。则不同分位数下的住院时间分布差异可表示为:

$$y_\tau^{UE} - y_\tau^N = \{y_\tau^{UE} - y_\tau^*\} + \{y_\tau^* - y_\tau^N\} \quad (7)$$

等式右边第一项由于研究问题的不同在工资性别差异分解中被称为“回报影响(return effect)”<sup>[23]</sup>或“性别歧视”<sup>[24]</sup>,在本研究中,它表示

由于参保方式不同所导致的患者住院时间差异部分,即“医保影响”;等式右边第二项则是“变量影响(covariate effect)”,它表示不同分位数下由于随机抽样的样本变量分布不同所导致的住院时间差异部分。

## 3 结果

### 3.1 描述性分析

本研究样本量大小为 36 329 人,住院时间均值为 11.86 天,中位数为 10 天,各变量的描述性统计结果详见表 1。对住院时间分布使用 k-s 法进行正态性检验,  $P < 0.001$ ,故住院时间不服从正态分布。对住院时间取对数后,按不同参保方式,绘制人口分布密度图(图 1)。可见城镇职工基本医疗保险、城镇居民基本医疗保险和新型农村合作医疗患者分布较集中且形状类似,但是新农合患者曲线最靠左,城职保患者的曲线最靠右,说明新农合患者的住院时间总体偏低,城职保患者的住院时间总体偏高,城居保患者居中;对于全公费患者,其人口密度最高点对应的住院时间和前三种相近,但是高住院时间的人口密度明显高于前三种,拖尾情况明显。

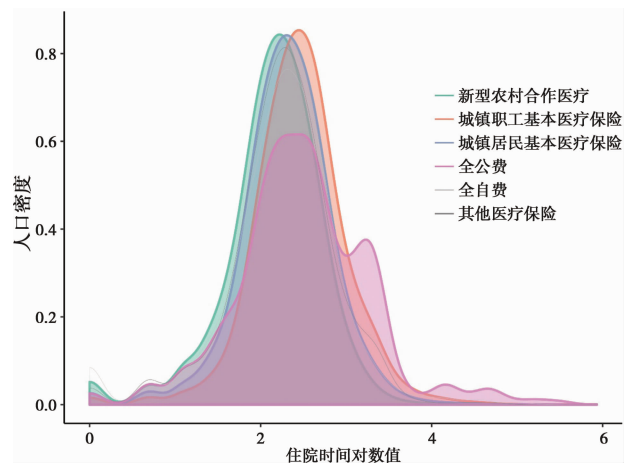


图 1 各参保方式患者住院时间样本分布

### 3.2 回归分析

基于式(6),建立住院时间与参保方式及各协变量的分位数回归,各分类变量分别以新型农村合作医疗、男性、疾病不严重、医嘱离院、未手术、未输血、未感染、甲等医院、综合医院和公立医院为对照,回归系数见表 2。其中模型(1)是只纳入了参保方式的简单回归模型,模型(2)是同时纳入了控制变量的全模型。中位数回归( $\tau = 0.5$ )的结果显示,参保方式效应值最大的是全公费(1.480),其次是城镇职工基

本医疗保险(1.178),表示在控制其他因素的情况下,全公费患者和城镇职工基本医疗保险比新农合患者在0.5分位数下住院时间多1.480天和1.178天。随着分位数的增加,各参保方式的系数均呈现上升趋势。而在各个分位数下比较模型(1)和模型

(2),可以发现参保方式系数符号相同,大小相近,随分位数增加系数增加趋势相同,说明模型比较稳健;此外,各分位数下模型(1)的系数普遍略大于模型(2)的系数,说明参保方式可能与各控制变量存在一定的正相关关系。

表2 分位数回归系数表

变量	QR estimates <sup>a</sup>										OLS <sup>b</sup>	
	0.10		0.25		0.50		0.75		0.90		(1)	(2)
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
参保方式(对照组:新型农村合作医疗)												
城镇职工基本医疗保险	3.000 <sup>***</sup>	0.832 <sup>***</sup>	2.000 <sup>***</sup>	0.853 <sup>***</sup>	3.000 <sup>***</sup>	1.178 <sup>***</sup>	3.000 <sup>***</sup>	1.618 <sup>***</sup>	6.000 <sup>***</sup>	2.719 <sup>***</sup>	3.805 <sup>***</sup>	2.197 <sup>***</sup>
城镇居民基本医疗保险	2.000 <sup>***</sup>	0.582 <sup>***</sup>	<0.000 1	0.520 <sup>***</sup>	<0.000 1	0.670 <sup>***</sup>	2.000 <sup>***</sup>	1.095 <sup>***</sup>	2.000 <sup>***</sup>	1.639 <sup>***</sup>	1.570 <sup>***</sup>	1.232 <sup>***</sup>
全公费	1.000 <sup>***</sup>	-0.032	1.000 <sup>***</sup>	0.382	3.000 <sup>***</sup>	1.480 <sup>***</sup>	7.000 <sup>***</sup>	4.085 <sup>***</sup>	13.000 <sup>***</sup>	7.781 <sup>***</sup>	7.455 <sup>***</sup>	5.615 <sup>***</sup>
全自费	<0.000 1	-0.220	<0.000 1	0.088	1.000 <sup>***</sup>	0.451 <sup>***</sup>	2.000 <sup>***</sup>	0.849 <sup>***</sup>	2.000 <sup>***</sup>	1.191 <sup>***</sup>	1.805 <sup>***</sup>	1.138 <sup>***</sup>
其他医疗保险	1.000 <sup>***</sup>	0.133	<0.001	0.314	1.000 <sup>***</sup>	0.823 <sup>***</sup>	2.000 <sup>***</sup>	1.193 <sup>***</sup>	3.000	1.671 <sup>***</sup>	1.490 <sup>***</sup>	0.915 <sup>***</sup>
女性	—	0.006	—	-0.088	—	-0.227 <sup>***</sup>	—	-0.496 <sup>***</sup>	—	-0.924 <sup>***</sup>	—	-0.397 <sup>***</sup>
年龄	—	0.006	—	0.010 <sup>***</sup>	—	0.019 <sup>***</sup>	—	0.027 <sup>***</sup>	—	0.038 <sup>***</sup>	—	0.034 <sup>***</sup>
疾病严重	—	0.128	—	0.608 <sup>***</sup>	—	1.023 <sup>***</sup>	—	1.333 <sup>***</sup>	—	2.061 <sup>***</sup>	—	1.197 <sup>***</sup>
离院方式(对照组:医嘱离院)												
非医嘱离院	—	-3.285 <sup>***</sup>	—	-3.020 <sup>***</sup>	—	-2.753 <sup>***</sup>	—	-2.608 <sup>***</sup>	—	-2.278 <sup>***</sup>	—	-2.941 <sup>***</sup>
死亡	—	-5.239 <sup>***</sup>	—	-5.520 <sup>***</sup>	—	-3.792 <sup>***</sup>	—	-0.019	—	4.472 <sup>*</sup>	—	-1.469 <sup>***</sup>
其他	—	-3.605 <sup>***</sup>	—	-2.902 <sup>***</sup>	—	-1.891 <sup>***</sup>	—	-0.697	—	0.743	—	-1.059 <sup>***</sup>
手术	—	0.737 <sup>***</sup>	—	1.221 <sup>***</sup>	—	2.227 <sup>***</sup>	—	3.806 <sup>***</sup>	—	6.491 <sup>***</sup>	—	3.293 <sup>***</sup>
输血	—	0.122	—	0.755 <sup>***</sup>	—	1.673 <sup>***</sup>	—	3.123 <sup>***</sup>	—	7.209 <sup>***</sup>	—	3.803 <sup>***</sup>
感染	—	0.506	—	0.696 <sup>***</sup>	—	0.840 <sup>***</sup>	—	1.181 <sup>***</sup>	—	1.81 <sup>***</sup>	—	1.078 <sup>***</sup>
乙等医院	—	-0.171	—	-0.108	—	0.075	—	0.430 <sup>***</sup>	—	1.279 <sup>***</sup>	—	0.554 <sup>***</sup>
床位数	—	<0.000 1	—	<0.000 1	—	<0.000 1	—	<0.000 1	—	<0.000 1	—	-0.000 1
机构类别(对照组:综合医院)												
精神病医院	—	0.021	—	0.015	—	0.242	—	2.047 <sup>***</sup>	—	6.381 <sup>***</sup>	—	4.529 <sup>***</sup>
传染病医院	—	0.228	—	1.510 <sup>***</sup>	—	1.865 <sup>***</sup>	—	4.016 <sup>***</sup>	—	6.08 <sup>***</sup>	—	5.057 <sup>***</sup>
职业病医院	—	1.538 <sup>***</sup>	—	2.176 <sup>***</sup>	—	1.510	—	0.873	—	0.129	—	0.887
非公立医院	—	0.354	—	0.721	—	0.404	—	1.582	—	3.691 <sup>***</sup>	—	0.163
职业	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
位置	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

注:<sup>a</sup>分位数回归估计值,标准误估计使用 bootstrap 法

<sup>b</sup>普通最小二乘估计值

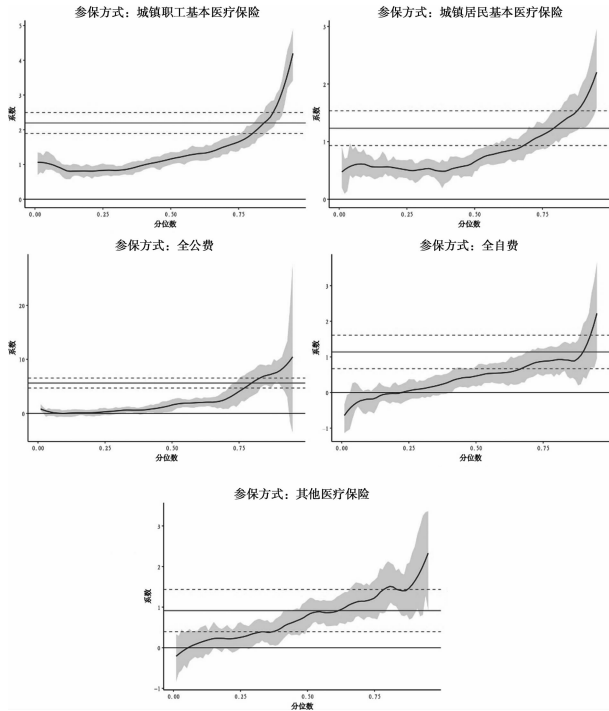
\*, \*\*, \*\*\* 分别表示  $P < 0.05$ 、 $P < 0.01$ 、 $P < 0.001$

通过建立各分位数下回归模型,得到各变量在各分位数下的系数变化情况,见图2(截距项被省略)。可以看出,除了性别为女性的曲线呈明显下降趋势,其他各变量的估计值随着分位数的增大,均呈现不同程度的上升趋势。参保方式中,全公费、全自费和其他医疗保险,在低分位数时估计没有统计学意义,而在中、高分位数有统计学意义;在0~0.5分位数下,各参保方式估计值上升速度较慢,曲线平坦,而在0.5~1分位数下,各估计值上升速度快,曲线陡峭。此外,OLS估计值普遍与条件分位数为0.8附近的估计值相近,体现出了OLS估计不能全面稳定的描述参保方式的效应的变化情况,而在本研究中高估了其效应值。

通过分位数回归模型,还能计算出调整了控制变量后不同参保方式在不同分位数下的住院时间的估计值(图3)。可见各参保方式患者住院时间确实存在差异,全公费的住院时间最长,且在0.5~0.75分位数间与其他各参保方式差距最大;城镇职工基本医疗保险、城镇居民基本医疗保险和新型农村合作医疗的曲线趋势类似,但是新型农村合作医疗的住院时间最少,甚至比全自费和其他医疗保险患者的住院天数更少。各参保方式住院时间随分位数上升速度特点为,0~0.5分位数间上升速度慢,0.5~1分位数间上升速度快,与系数估计值变化的结论相符。

从图3可以看出,新型农村合作医疗患者的住院

在 70% ~ 80% 左右,受样本影响度在 20% 左右,而城镇职工基本医疗保险平均医保影响度在 60% 左右,可见城镇职工基本医疗保险与新型农村合作医疗的差异大部分也是由医保不同造成的,但是相比全公费和城镇居民基本医疗保险,医保的影响度稍低。



注:限于篇幅有限,此图以参保方式为例

图 2 变量系数估计值变化图

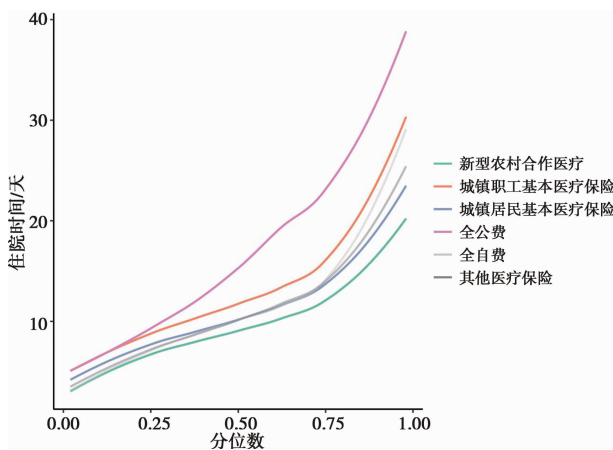


图 3 各参保方式患者住院时间在不同分位数下的分布

时间最短,与其他参保方式存在差异,那么,这种差异是由于不同参保方式的样本不同引起还是由于参保方式不同引起的,为了解决这个问题,分别对新型农村合作医疗与城镇职工基本医疗保险、城镇居民基本医疗保险和全公费医疗的差异进行了分解。通过构建反事实分布,将差异分解成两部分,即“医保差异”和“变量差异”(表 3 - 表 5)。各参保方式随分位数增加,总差异增加,医保影响占总差异百分比即医保影响度也在增加,而变量差异增加较少或趋于稳定,故变量影响度反而减少,而在极高分位数下,样本之间差异明显增大,变量影响占总差异比例相比中高分位数有所升高。全公费医疗和城镇居民基本医疗保险平均医保影响度

表 3 新型农村合作医疗和全公费医疗住院时间分解

分位数	总差异 (95% CI <sup>a</sup> )	医保影响 (95% CI <sup>a</sup> )	医保影响度 (%)	变量影响 (95% CI <sup>a</sup> )	变量影响度 (%)
0.1	2.33 (1.43 - 3.32)	1.38 (0.38 - 2.52)	59.36	0.95 (0.40 - 1.45)	40.64
0.2	2.97 (1.82 - 4.47)	2.07 (0.86 - 3.61)	69.69	0.90 (0.37 - 1.34)	30.31
0.3	3.94 (2.56 - 6.05)	3.08 (1.68 - 5.24)	78.23	0.86 (0.42 - 1.34)	21.77
0.4	5.13 (3.29 - 7.98)	4.19 (2.34 - 7.2)	81.76	0.94 (0.47 - 1.57)	18.24
0.5	6.49 (3.29 - 7.98)	5.35 (2.84 - 8.73)	82.50	1.14 (0.47 - 1.57)	17.50
0.6	8.19 (5.16 - 11.46)	6.77 (3.59 - 10.03)	82.68	1.42 (0.66 - 2.78)	17.32
0.7	9.80 (6.31 - 13.11)	8.01 (4.26 - 11.62)	81.80	1.78 (0.83 - 3.55)	18.20
0.8	10.97 (7.67 - 14.69)	8.61 (4.18 - 12.76)	78.54	2.35 (0.88 - 4.38)	21.46
0.9	13.59 (8.8 - 23.17)	9.60 (1.77 - 19.9)	70.69	9.98 (1.14 - 9.34)	29.31

注: <sup>a</sup> CI (Confidence Interval) 置信区间,通过 bootstrap 计算

表 4 新型农村合作医疗和城镇职工基本医疗保险住院时间分解

分位数	总差异 (95% CI <sup>a</sup> )	医保影响 (95% CI <sup>a</sup> )	医保影响度 (%)	变量影响 (95% CI <sup>a</sup> )	变量影响度 (%)
0.1	1.91 (1.74 - 2.08)	0.61 (0.05 - 1.19)	31.8	1.30 (0.71 - 1.86)	68.2
0.2	1.97 (1.82 - 2.11)	0.77 (0.24 - 1.3)	39.2	1.20 (0.66 - 1.69)	60.8
0.3	2.14 (2.01 - 2.27)	1.02 (0.45 - 1.51)	47.5	1.12 (0.61 - 1.66)	52.5
0.4	2.4 (2.25 - 2.54)	1.21 (0.47 - 1.75)	50.3	1.19 (0.66 - 1.91)	49.7
0.5	2.71 (2.55 - 2.86)	1.27 (0.51 - 1.89)	46.7	1.44 (0.84 - 2.22)	53.3
0.6	3.06 (2.88 - 3.24)	1.44 (0.6 - 2.27)	46.9	1.63 (0.8 - 2.42)	53.1
0.7	3.56 (3.34 - 3.78)	1.70 (0.11 - 2.59)	47.6	1.86 (1 - 3.4)	52.4
0.8	4.4 (4.06 - 4.77)	2.19 (0.19 - 3.52)	49.8	2.21 (0.91 - 4.2)	50.2
0.9	6.57 (5.97 - 7.21)	2.82 (-3.57 - 5.78)	43.0	3.75 (0.98 - 10.04)	57

注: <sup>a</sup> CI (Confidence Interval) 置信区间,通过 bootstrap 计算

表5 新型农村合作医疗和城镇居民基本医疗保险住院时间分解

分位数	总差异 (95% CI <sup>a</sup> )	医保影响 (95% CI <sup>a</sup> )	医保影 响度(%)	变量影响 (95% CI <sup>a</sup> )	变量影 响度(%)
0.1	1.06 (0.87 - 1.26)	0.67 (0.38 - 1.01)	63.01	0.39 (0.15 - 0.62)	36.99
0.2	1.02 (0.87 - 1.19)	0.63 (0.38 - 0.86)	61.17	0.40 (0.19 - 0.59)	38.83
0.3	0.99 (0.84 - 1.14)	0.65 (0.42 - 0.85)	65.62	0.34 (0.18 - 0.54)	34.38
0.4	1.03 (0.87 - 1.20)	0.69 (0.43 - 0.93)	67.08	0.34 (0.14 - 0.54)	32.92
0.5	1.13 (0.94 - 1.32)	0.73 (0.46 - 1.00)	64.72	0.40 (0.2 - 0.62)	35.28
0.6	1.30 (1.09 - 1.53)	0.85 (0.52 - 1.17)	65.49	0.45 (0.21 - 0.71)	34.51
0.7	1.52 (1.29 - 1.80)	1.02 (0.60 - 1.40)	67.42	0.49 (0.20 - 0.87)	32.58
0.8	1.72 (1.43 - 2.05)	1.23 (0.66 - 1.72)	71.22	0.50 (0.12 - 0.97)	28.78
0.9	2.39 (1.93 - 2.86)	1.59 (0.13 - 2.43)	66.63	0.80 (0.14 - 2.14)	33.37

注:<sup>a</sup> CI(Confidence Interval)置信区间,通过 bootstrap 计算

#### 4 讨论

通过对住院时间建立各影响因素的分位数回归模型,可以看出对住院时间影响较大的几个因素是参保方式、离院方式死亡、是否手术、是否输血,这些因素的效应值较大,而且随着分位数增加的变化程度也较大。在其他因素一定的情况下,进行了手术的患者和进行了输血的患者的住院天数比没有进行的在不同分位数下高1~6天,效应呈递增趋势,因为当COPD患者需要进行手术或者输血时往往是由于急性感染等出现了并发症,这时患者需要更多的医疗服务利用以及更长的住院治疗时间,所以它们是住院时间的重要影响因素。对于离院方式死亡,其系数呈先下降后上升的趋势,符号在低分位数时为负,在高分位数时为正。说明在低分位数时,即住院时间较短时,死亡患者比医嘱离院患者有更少的住院天数,因为住院时间较低的死亡患者往往病情较急,还没有利用足够的医疗服务就发生了死亡结局,故住院天数更短;而在高分位数时,即住院时间较长时,死亡患者比医嘱离院的患者有更多的住院天数,因为作为慢性病,COPD患者在长期住院过程中由于一直没有好转导致死亡结局,所以死亡患者比好转后医嘱离院患者会接受更多的治疗和服务,故住院时间会更长。由于离院方式是一个重要的住院时间影响因

素,其中涉及患者治愈、好转、未治愈和死亡等状态,所以有研究也建议可用生存分析的方法研究住院时间。<sup>[25]</sup>

参保方式中,城镇职工基本医疗保险和城镇居民基本医疗保险在各分位数下均有统计学意义,说明不管住院时间长或短,城职保和城居保患者都比新农合患者利用更多医疗服务;全公费、全自费和其他医疗保险在低分位数下没有意义,在高分位数下有统计学意义,说明在住院时间较短时,全公费、全自费、其他医疗保险和新农合患者卫生资源利用情况相同,没有差异,但在住院时间较长时,新农合患者的住院服务利用度较低。

在新农合和全公费、城镇职工医保、城镇居民医保住院时间差异上,随着分位数增加,新农合与其他参保方式的差异呈递增趋势且有统计学意义,分解出的“医保影响”占总差异的大部分,说明排除了各样本变量分布差异的影响后,医保差异仍然是造成住院时间差异的主要因素。在极高分位数下( $\tau = 0.9$ ),“变量影响度”有一定程度的上升,说明作为样本分布的极端边界,住院时间极长的患者间的样本差异是较大的。

在调整了其他因素后全公费住院时间最长,新农合最短,在中高分位数下,新农合甚至比全自费、医疗救助和其他社会保险更短。其中可能的原因包括患者存在医疗服务利用不足或过度利用,医院存在提供服务不足或过度提供,以及不同医疗保险的主办方监管力度不同。如果以全自费的患者为对照,假设医院考虑到患者没有医保报销,所以不会给全自费患者进行过量的诊疗服务。在控制了疾病相关因素后,全公费、城职保、城居保患者的住院天数仍然高于全自费,而患者很难自发延长自己住院时间,说明医院存在提供过度医疗服务的情况,而且对不同参保方式患者提供的过度医疗量不同,在住院时间较短患者中主要对城镇职工医保和城镇居民医保患者提供过度医疗,在住院时间较长的患者中主要对全公费和城职保患者提供过度医疗服务。而新农合患者的住院时间短于全自费,可能是由于新农合患者住院费用通常比城职保和全公费低的,故医院为了最大化经济效益,选择将新农合患者的床位转移给住院费用更高的全自费和城职保患者。这是一个较保守的假设,因为如果假设不成立,即医院会向全自费患者提供过度医疗,那么向全公费、城职保等患者提供过度医疗的情况将会更严重。除此之

外,新农合的基金池压力大,城镇居民医保和城镇职工医保等保险的基金池压力较小,使得各医保主办方对医院的监管力度不同,从而也可能导致新农合与各医疗保险住院天数产生差异。

目前,医疗保险费用结算的方式大多采用的是总额控制下的按医疗服务项目付费的后付制。由于是后付制度,如果缺乏有效的制约机制,可能极大地刺激过度消费,诱使医院利用医患双方信息不对称提供过度医疗服务,对平均住院日的缩短产生负面影响。目前国外已有一些地区的医院尝试通过罚金的方式制约这种过度消费,但是 Tor Helge 等通过一项自然实验研究发现,使用罚金反而会增加患者的住院时间。<sup>[26]</sup>因此,为了消除这种过度消费,应当实施支付方式改革,建立总额预付制度。

本研究的分析方法是基于分位数回归模型的实证分析,尽管是一种相关关系的分析方法,不能回答因果关系,但结合本研究的具体内容,结果也有一定的因果含义。参保状态可以影响住院日,但一个人的参保状态很难由住院日来决定。此外,本研究也没有较明显的内生性问题。本研究的不足之处在于,患者教育程度和收入可能是住院时间较为重要的两个影响因素,但是由于本研究数据的限制,无法获取,可能会对参保方式系数的估计产生遗漏变量偏误,但是本研究纳入了地区哑变量,可以一定程度弥补收入等经济因素带来的问题。由于病案首页上医疗付款方式为单选,代表报销所用的参保方式,但是实际上一名患者除了参保基本医疗保险以外还可能参保其他非基本医疗保险,因此,医疗救助、商业医院保险对象往往会和新型农村合作医疗、城镇居民基本医疗保险和城镇职工基本医疗保险存在一定的重叠,而从现有数据中无法识别,但是本研究中参保方式关注的正是患者报销用的参保方式,可能患者参保了某种基本医疗保险但是报销使用的是其他医疗保险,而这个报销用的保险才是影响患者住院时间的因素。此外,本研究只纳入了四川省全部三级医院第四季度的患者,不同参保方式在不同等级医院和不同季度的关系和住院时间差异上是否有所不同,需要更进一步的研究。

**作者声明本文无实际或潜在的利益冲突。**

#### 参 考 文 献

[1] 慢性阻塞性肺疾病(COPD)诊治规范(草案)[J]. 中华

结核和呼吸杂志, 1997(4): 1-3.

- [2] 唐文芳, 刘日辉, 于雅琴, 等. 2000—2014 年中国 40 岁以上成人慢性阻塞性肺疾病患病率的 Meta 分析[J]. 吉林大学学报(医学版), 2015, 41(5): 961-968.
- [3] Collaborators M C O D. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013 [J]. *Lancet*, 2015, 385(9963): 117-171.
- [4] 王玲, 王承珠, 缙文海. 某院 316 例超长住院日病例统计分析[J]. 中国卫生统计, 2015, 32(2): 367-368.
- [5] 刘延徽. 2236 例超长住院患者住院时间的影响因素分析[J]. 中国肿瘤, 2016, 25(12): 1022-1024.
- [6] 王霞, 郭秀娥, 徐勇勇, 等. 住院手术患者无效住院日的存在状况及其影响因素的分析[J]. 医学争鸣, 2001, 22(6): 557-559.
- [7] 吕艳伟, 张国英, 邵文玲. 股骨骨折病人住院日水平及影响因素分析[J]. 中国卫生统计, 2010, 27(6): 617-618.
- [8] 王箐. 医疗保险对医疗质量影响的研究[J]. 经济与管理研究, 2012, (10): 24-31.
- [9] Cournane S, Conway R, Creagh D, et al. Radiology imaging delays as independent predictors of length of hospital stay for emergency medical admissions [J]. *Clin Radiol*, 2016, 71(9): 912-918.
- [10] Jawitz O K, Boffa D J, Detterbeck F C, et al. Estimating the Annual Incremental Cost of Several Complications Following Pulmonary Lobectomy [J]. *Semin Thorac Cardiovasc Surg*, 2016, 28(2): 531-540.
- [11] Zhang J H, Yuan J, Wang T. Direct cost of dengue hospitalization in Zhongshan, China: Associations with demographics, virus types and hospital accreditation [J]. *PLoS Negl Trop Dis*, 2017, 11(8): e0005784.
- [12] Siciliani L, Sivey P, Street A. Differences in length of stay for hip replacement between public hospitals, specialised treatment centres and private providers: selection or efficiency? [J]. *Health Econ*, 2013, 22(2): 234-242.
- [13] De La Garza-Ramos R, Goodwin C R, Abu-Bonsrah N, et al. Prolonged length of stay after posterior surgery for cervical spondylotic myelopathy in patients over 65 years of age [J]. *J Clin Neurosci*, 2016, 31: 137-141.
- [14] Park K U, Rubinfeld I, Hodari A, et al. Prolonged Length of Stay after Esophageal Resection: Identifying Drivers of Increased Length of Stay Using the NSQIP Database [J]. *J Am Coll Surg*, 2016, 223(2): 286-290.
- [15] Deluzio M R, Keshava H B, Wang Z, et al. A model for predicting prolonged length of stay in patients undergoing anatomical lung resection: a National Surgical Quality Im-



provement Program (NSQIP) database study[J]. *Interact Cardiovasc Thorac Surg*, 2016, 23(2): 208-215.

[16] 杨天桂, 石应康, 杨洋. 影响缩短平均住院日因素和措施的循证研究[J]. *中国医院*, 2008, 12(12): 51-54.

[17] 寇长贵, 刘洁楠, 俞琼, 等. 精神分裂症患者住院时间影响因素分析[J]. *中国公共卫生*, 2009, 25(3): 302-303.

[18] Wooldridge J M. *Introductory econometrics: a modern approach*[M]. 2005.

[19] Koenker R, Bassett G. *Regression Quantiles*[J]. *Econometrica*, 1978, 46(1): 33-50.

[20] 季莘, 陈峰. 百分位数回归及其应用[J]. *中国卫生统计*, 1998(6): 9-11.

[21] 李育安. 分位数回归及应用简介[J]. *统计与信息论坛*, 2006, 21(3): 35-38.

[22] 郭继强, 姜俪, 陆利丽. 工资差异分解方法述评[J].

*经济学: 季刊*, 2011, 10(2): 363-414.

[23] 段景辉, 陈建宝. 我国城乡家庭收入差异影响因素的分位数回归解析[J]. *经济学家*, 2009(9): 46-53.

[24] 陈建宝, 段景辉. 中国性别工资差异的分位数回归分析[J]. *数量经济技术经济研究*, 2009(10): 87-97.

[25] 时俊新, 张静, 蒋春舫等. 生存分析方法在道路交通事故致颅脑损伤住院时间估计中的应用[J]. *中国卫生统计*, 2008, 25(6): 573-576.

[26] Holmås T H, Kjerstad E, Lurås H, et al. Does monetary punishment crowd out pro-social motivation? A natural experiment on hospital length of stay[J]. *Journal of Economic Behavior & Organization*, 2010, 75(2): 261-267.

[收稿日期: 2018-03-06 修回日期: 2018-06-06]  
(编辑 薛云)

· 信息动态 ·

## 欢迎订阅 2019 年《中国卫生政策研究》杂志

《中国卫生政策研究》杂志是国家卫生和计划生育委员会主管, 中国医学科学院主办, 医学信息研究所和卫生政策与管理研究中心承办的卫生政策与管理专业学术期刊, 国际标准连续出版物号为 ISSN 1674-2982, 国内统一刊号为 CN 11-5694/R, 本刊为中文核心期刊、中国科学引文数据库(CSCD)核心期刊、中国科技核心期刊(中国科技论文统计源期刊)、RCCSE 中国核心学术期刊(A)、《中国人文社会科学期刊评价报告(AMI)》引文数据库期刊、人大复印报刊资料数据库重要转载来源期刊。

杂志以“传播政策、研究政策、服务决策”为办刊方针, 及时报道卫生政策研究最新成果和卫生改革发展新鲜经验, 促进卫生政策研究成果的传播利用及卫生政策研究者与决策者的交流合作, 提高卫生政策研究理论水平和实践能力, 为政府科学决策、改进卫生绩效和促进卫生事业发展提供重要学术支撑。主要适合各级卫生行政部门和卫生事业单位管

理者、卫生政策与管理相关领域的专家学者和实践者、高等院校相关专业的师生等阅读。主要栏目有: 专题研究、医改进展、卫生服务研究、医疗保障、药物政策、社区卫生、农村卫生、公共卫生、医院管理、全球卫生、卫生人力、卫生法制、理论探讨、经验借鉴、书评等。

杂志为月刊, 每月 25 日出版, 国内外公开发行, 大 16 开本, 进口高级铜版纸彩封印刷, 定价 20 元/册, 全年 240 元(含邮资)。

全国各地邮局均可订阅, 邮发代号 80-955, 也可向编辑部直接订阅。

地址: 北京市朝阳区雅宝路 3 号中国医学科学院医学信息研究所《中国卫生政策研究》编辑部

邮编: 100020

E-mail: cjhp@imicams.ac.cn

联系人: 薛云

电话: 010-52328696、52328697