

指标选择对医院效率评价的影响

——以 2010 年省级数据 DEA 模型为例

石义全^{1*} 钱振华² 成刚³

1. 山东省烟台市疾病预防控制中心 山东烟台 264003
2. 北京科技大学 北京 100083
3. 北京大学中国卫生发展研究中心 北京 100191

【摘要】目的:探讨在应用 DEA 方法评价医院效率时,指标选择对分析结果的影响。方法:(1)以 31 个省 2010 年的数据为例,建立一系列包含不同投入产出指标数量的数据包络分析模型,观察指标选择对分析结果的影响;(2)通过实例说明专题地图在展示效率地域分布特征中的作用。结果:随着投入产出指标数量的增加,DEA 分析结果的区分度逐渐降低,有效决策单元数量由 5 个增加到 17 个,默认有效单元的数量由 3 个增加到 5 个,效率均值由 0.867 上升到 1.056。各地区医院效率分布大致呈现南、中、北的分布特征。结论:投入产出指标的数量严重影响 DEA 模型的分析结果,在数据支持的情况下,应该通过比较多种投入产出指标的组合以检验 DEA 分析结论的可靠性。专题地图在展示效率地域分布特征时具有简单、直观的优势,是发现效率影响因素和进行实证研究的基础。

【关键词】数据包络分析;效率;医院;指标选择;专题地图

中图分类号:R197 文献标识码:A doi: 10.3969/j.issn.1674-2982.2012.03.013

Impact of indicator selection on hospital efficiency assessment using data envelopment analysis:

An empirical application with 31 provincial data in 2010

SHI Yi-quan¹, QIAN Zhen-hua², CHENG Gang³

1. Yantai Center for Disease Prevention and Control, Shandong Yantai 264003, China
2. University of Science and Technology Beijing, Beijing 100083, China
3. China Center for Health Development Studies, Peking University, Beijing 100191, China

【Abstract】 Objective: To explore the impact of indicator selection on hospital efficiency assessment using data envelopment analysis. Methods: A series of DEA models are constructed using different number of input and output indicators, and observes the impact of the selected number of inputs and outputs on the results. The hospital data of 31 provinces in 2010 are used. A thematic map is used to show the geographical distribution of the efficiency analysis results. Results: The discriminating power of DEA results is remarkably reduced in line with the number of inputs and outputs is increasing. For instance, the effective decision-making units are increased from 5 to 17 as well as the default effective decision-making units increased from 3 to 5 respectively. In addition, the scale mean efficiency is increased from 0.867 to 1.056. The geographical distribution of efficiency scores shows decreasing tendency from south to north. Conclusion: The DEA result is highly dependent on the number of inputs and outputs. Multiple comparison using different combinations of inputs and outputs should be made to validate the conclusions, if data are available. A thematic map is a useful tool to estimate influencing factors and solid empirical evidence as regarding to show the geographical distribution of hospital efficiency.

【Key words】 Data Envelopment Analysis; Efficiency; Hospital; Indicator Selection; Thematic map

* 基金项目:中国博士后科学基金(20090450259)

作者简介:石义全,男(1974年—),主管医师,主要研究方向为卫生政策。E-mail: leonstone@163.com

通讯作者:成刚。E-mail: chenggang@bjmu.edu.cn

效率评价是卫生体系研究的重要内容,医疗服务效率是卫生体系效率评价的一个重要环节。医院是医疗服务提供的主要场所,根据2011年《中国卫生统计年鉴》的数据,医院提供的门诊服务数量约占所有医疗机构(不包括村卫生室)门诊服务数量的1/2,医院提供的住院服务数量占有所有医疗机构住院服务数量的60%以上。在卫生效率评价的国内外文献中,数量最多的是对医院效率的评价。由于医院是多投入、多产出的生产单位,因此,医院效率评价最常用的方法是能够处理多指标的两类评价方法:数据包络分析(Data envelopment analysis, DEA)和随机前沿分析(Stochastic frontier analysis, SFA)^[1-2]。

DEA是非参数技术效率评价方法,对于分析多投入、多产出的问题具有明显优势,自1978年Charnes等人创立该方法以来^[3],DEA在各领域得到了广泛的应用^[4-5]。1981年Sherman首先将DEA应用于医院效率评价^[6],此后在卫生领域的应用越来越多^[7-12]。DEA虽然具有特殊的优势,但也有明显的弱点^[1-2]。首先,投入产出指标数量和样本数量严重影响DEA分析结果^[13-14];对于同一样本,投入产出指标数量增加会使有效单元比例和平均效率值增加;对于相同的投入产出指标,在样本数量较少时,减少样本数量也会使有效单元比例和平均效率值增加。其次,在可变规模报酬(Variable returns to scale, VRS)模型中,当样本数量与指标数量的比值较小时,默认有效会成为影响分析结果的突出问题。默认有效是指被评价的决策单元由于没有相似的参比对象而自动被判定为有效。在DEA模型中,选择恰当的投入产出指标是得出正确结论的前提。

近年来,DEA方法在国内卫生领域中的应用越来越多,但是在应用中往往忽略了DEA的上述特点,从而得出不恰当的结论。常见问题包括:(1)指标选择不当,例如在健康指标中用总人口死亡率代替婴儿死亡率^[15-17]; (2)投入产出指标数量过多,在DEA效率区分度很低的情况下,得出总体效率较高的结论^[18]; (3)忽略了对默认有效问题的分析,这是一个普遍存在的问题。另外,在效率结果的分析中,往往习惯性地划分为东、中、西部^[16,19,20],而效率高低的真实分布未必如此,这样就会影响对潜在影响因素的发现。本文的目的是针对上述问题,通过实证分析探讨在应用DEA方法评价医院效率时,投入产出指标数量对分析结果的影响,并通过实例说明专题地图在分析效率地域分布特征中的作用。

1 资料与方法

1.1 资料来源

分析数据来自2011年《中国卫生统计年鉴》,包括:2010年各地区卫生人员数、2010年各地区医疗机构床位数、2010年各地区医疗机构门诊服务情况以及2010年各地区医疗机构住院服务情况。

1.2 研究方法

分析数据包含了31个决策单元(省、直辖市、自治区),依据文献中关于样本数量与指标数量关系的规则^[13-14,21],投入和产出的指标数量最多可以达到10个。为了比较指标数量对效率分析结果的影响,本文采用不同的指标组合,共形成8个模型,指标数量分别为3~10个(表1)。

表1 不同投入和产出指标组合

模型	投入指标	产出指标	投入指标数量	产出指标数量	指标总数
模型1	卫生人员数	诊疗人次数、入院人数	1	2	3
模型2	床位数、卫生人员数	诊疗人次数、入院人数	2	2	4
模型3	床位数、卫生人员数	诊疗人次数、入院人数、住院手术人次数	2	3	5
模型4	床位数、卫生人员数	诊疗人次数、入院人数、住院手术人次数、危重病人抢救人次数	2	4	6
模型5	床位数、卫生人员数	诊疗人次数、健康检查人数、入院人数、住院手术人次数、危重病人抢救人次数	2	5	7
模型6	床位数、卫生人员数	诊疗人次数、留观病例数、健康检查人数、入院人数、住院手术人次数、危重病人抢救人次数	2	6	8
模型7	床位数、执业(助理)医师数、注册护士数	诊疗人次数、留观病例数、健康检查人数、入院人数、住院手术人次数、危重病人抢救人次数	3	6	9
模型8	床位数、执业(助理)医师数、注册护士数、其它人员数	诊疗人次数、留观病例数、健康检查人数、入院人数、住院手术人次数、危重病人抢救人次数	4	6	10

指标之间的高度相关(共线性)并不会影响 DEA 模型的结果,但为了观察指标间相关程度在指标数量与分析结果关系中的作用,表 2 列出了投入指标之间的相关系数,表 3 列出了产出指标之间的相关系数。

表 2 投入指标之间的相关系数矩阵

	床位数	卫生 人员数	执业(助理) 医师数	注册 护士数	其它 人员数
床位数	1.000				
卫生人员数	0.979	1.000			
执业(助理)医师数	0.982	0.991	1.000		
注册护士数	0.973	0.995	0.981	1.000	
其它人员数	0.962	0.993	0.974	0.982	1.000

表 3 产出指标之间的相关系数矩阵

	诊疗人 次数	留观病 例数	健康检 查人数	入院 人数	住院手术 人次	危重病人 抢救人次
诊疗人次	1.000					
留观病例数	0.966	1.000				
健康检查人数	0.441	0.488	1.000			
入院人数	0.786	0.846	0.619	1.000		
住院手术人次	0.933	0.963	0.587	0.944	1.000	
危重病人抢救人次	0.502	0.569	0.748	0.833	0.717	1.000

根据健康生产理论与实证分析,卫生体系的生产技术是规模报酬可变的,因此选择 VRS 模型,即 BCC 模型。^[22-23]对于有效的决策单元采用超效率值、被参比的次数和总权重进行综合评价(详见本期《卫生体系效率评价的概念框架与测量方法——兼论应用数据包络分析的方法学问题》)。模型方向选择产出导向。DEA 分析工具为 MaxDEA 软件。采用专题地图展示效率的地区分布特征,制图软件采用 ArcGIS。

2 结果

2.1 指标数量对分析结果的影响

从模型 1 到模型 8 投入产出指标的数量由 3 个逐步增加到 10 个,效率评价结果的区分度逐渐降低:有效单元数量由 5 个增加到 17 个;默认有效单元

的数量由 3 个增加到 5 个;最小效率值由 0.520 上升到 0.621,均值由 0.867 上升到 1.056(表 4)。

2.2 效率整体情况

效率整体情况可通过效率均值来体现,由模型 1 得出的效率均值最小,为 0.867,但模型 1 的投入指标仅包含了卫生人力一个方面,不够全面。模型 2 的投入指标包含人力和床位数两个方面,产出指标包含门诊和住院两个方面,这是医院效率评价文献中最常用的指标组合。模型 3 在模型 2 的基础上增加了住院手术数量,在一定程度上反映了病例的严重程度。模型 4—8 的结果中有效单元数量过多,区分度较低。因此,用模型 2 和模型 3 的结果判断整体效率情况更为合适。从模型 2 和模型 3 的效率均值(分别为 0.944 和 0.948)看,各省医院效率的整体差异不大。两个模型得出的效率最低的五个省份完全相同,其效率值也几乎完全相同(除 1 个省份相差 0.002 以外,其它 4 个省份相同),五个省份是山西、黑龙江、吉林、内蒙古和辽宁,其效率值分别为 0.61、0.69、0.69、0.71、0.73,提示这 5 个省份存在较大的效率提升空间。

2.3 对有效单元的综合分析

表 5 是被评价为有效的省份被参比的次数和总权重。限于篇幅,仅列出了模型 1—3 的结果,数据和完整的分析结果可从网上下载(http://www.maxdea.cn/Sharing/DEA_China2010.rar)。

被评价为有效的决策单元(包括默认有效的决策单元)作为标杆的重要性可通过被其它决策单元参比的次数和总权重来体现,被参比次数越多、总权重越大,其标杆意义越大。广东、山东和西藏是 3 个默认有效的省份,其中西藏在所有模型中的超效率值均为 1,不存在“超效率”,说明西藏被评价为有效完全是因为其投入指标数值太小,没有相似的省份进行参比,所以虽然西藏处于生产前沿,但是对其它省份没有参比的意义。在 3 个模型中,超效率值最大的均为广东。在模型 1 中,被参比次数最多、总权重最大的为云南。在模型 2 和模型 3 中,被参比次数最多、总权重最大的省份是福建省。综合分析提示,广东、福建、云南 3 省的标杆意义相对比较突出。

表4 各地区不同模型的效率值(超效率模型)

地区	模型1	模型2	模型3	模型4	模型5	模型6	模型7	模型8
安徽	0.859	0.945	0.945	0.945	0.945	0.945	0.943	0.956
北京	0.698	0.881	0.884	0.884	0.884	0.947	0.947	0.947
福建	0.973	1.066	1.066	1.066	1.066	1.066	1.046	1.100
甘肃	0.771	0.784	0.784	0.784	0.810	0.810	0.820	0.823
广东	1.643*	1.643*	1.676*	1.681*	1.878*	1.878*	1.878*	1.878*
广西	0.803	0.996	0.996	1.017	1.018	1.018	1.024	1.024
贵州	0.891	0.982	0.982	0.982	0.982	1.013	1.014	1.016
海南	0.665	0.921	0.921	0.921	0.921	0.921	0.921	0.921
河北	0.908	0.955	0.955	0.955	0.955	0.955	0.994	0.994
河南	0.931	0.941	0.941	1.027*	1.027*	1.027*	1.027*	1.027*
黑龙江	0.619	0.689	0.691	0.691	0.691	0.691	0.715	0.715
湖北	0.902	0.985	0.985	1.043	1.043	1.043	1.043	1.043
湖南	0.973	1.025	1.025	1.025	1.025	1.025	1.041	1.041
吉林	0.595	0.694	0.694	0.695	0.695	0.695	0.727	0.727
江苏	0.964	0.964	0.964	0.964	0.964	0.964	0.954	0.992
江西	0.758	0.982	0.982	0.983	0.983	1.022	1.022	1.022
辽宁	0.709	0.733	0.733	0.733	0.733	0.753	0.753	0.753
内蒙古	0.584	0.713	0.713	0.748	0.748	0.748	0.755	0.755
宁夏	0.803	0.901	0.901	0.926	0.926	0.938	0.939	0.939
青海	0.803	0.872	0.872	1.292	1.292	1.292	1.292	1.301
山东	1.110*	1.110*	1.110*	1.110*	1.110*	1.110*	1.075*	1.221*
山西	0.520	0.608	0.608	0.608	0.608	0.608	0.621	0.621
陕西	0.707	0.841	0.841	0.844	0.844	0.844	0.852	0.852
上海	1.168	1.168	1.168	1.168	1.168	1.168	1.192	1.238
四川	0.984	0.998	1.009	1.072*	1.072*	1.235*	1.234*	1.236*
天津	0.763	0.910	0.961	0.961	0.961	1.379	1.379	1.379
西藏	1.000*	1.000*	1.000*	1.000*	1.000*	1.000*	1.000*	1.000*
新疆	0.850	0.945	0.945	0.945	0.945	0.945	0.966	0.966
云南	1.186	1.186	1.186	1.569	1.569	1.652	1.770	1.770
浙江	0.958	0.958	0.958	0.958	0.958	0.958	0.912	1.051
重庆	0.770	0.860	0.904	0.904	0.904	1.291	1.429	1.429
最小值	0.520	0.608	0.608	0.608	0.608	0.608	0.621	0.621
均值	0.867	0.944	0.948	0.984	0.991	1.030	1.042	1.056
有效单元个数	5	7	8	12	12	16	16	17
默认有效个数	3	3	3	5	5	5	5	5

注:有效决策单元的效率值为超效率值(效率值大于或等于1);标注*号的单元为默认有效。

表5 有效的省份被其它省份参比的次数和总权重

地区	模型1		模型2		模型3	
	参比次数	总权重	参比次数	总权重	参比次数	总权重
福建	—	—	17	8.607	17	8.128
广东	11	1.324	14	2.099	14	2.264
广西	—	—	—	—	0	0
湖南	—	—	9	3.437	8	3.011
山东	13	3.928	5	1.325	4	0.965
上海	11	2.165	3	1.748	3	1.438
西藏	10	4.265	10	4.349	11	4.787
云南	25	14.318	12	2.435	12	2.406

2.4 医院效率的地域分布特征

相对而言,模型2和模型3最适宜分析整体效率情况。图1以模型2和模型3的分析结果为例,演示了专题地图在显示效率分布特征中的作用(由于两个模型得出效率值差异很小,所以在专题地图上的分布完全相同)。从专题地图可以发现,各地区医院效率大致呈现南、中、北的分布特征,效率相对最高的地区是西南、南部、东南及东部沿海,效率相对最低的地区是东北、北部,效率中等的地区是西北和中西部(图1)。

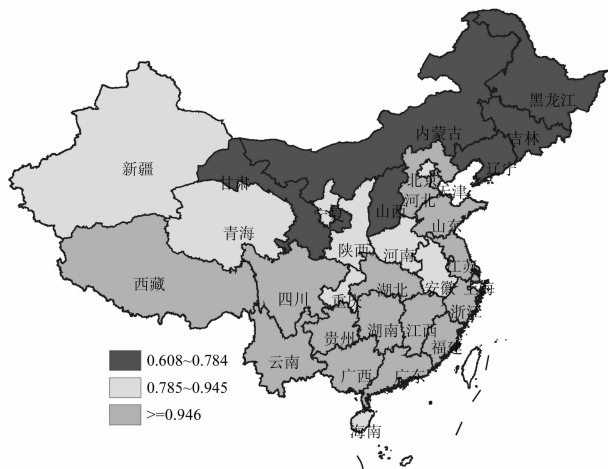


图1 医院效率的区域分布(模型2和模型3数据)

3 讨论

3.1 DEA 模型的指标选择

与参数方法相比,DEA 模型对样本数量的要求较低。^[1]一般遵循的原则是样本数量不应少于投入和产出指标数量的 3 倍。^[13-14,21]在实证研究中,一般情况下研究对象(即样本)是确定的,投入和产出指标的选择具有一定的弹性。依据上面的规则,投入和产出指标的数量要求不超过样本数量的 1/3,但这一原则是根据数据模拟结果和效率区分度的最低要求确定的。在实证研究中,在满足上述要求的情况下,实际区分度未必能够达到分析的需要。

DEA 模型中包含的投入和产出指标,应该存在以下关系:由投入指标可以生产出产出指标,产出指标是由投入指标生产出来的。但在卫生领域,投入和产出之间的关系并不像工厂的生产车间那样简单。以医院为例,投入方面主要是卫生人员和医疗设备,但除此以外,医院的基础设施建设、科研投入等也是投入指标;产出方面主要是门诊和住院服务,但除此以外,不同级别的医院还可能承担着医学教育、健康教育、传染病防治、对口支援贫困地区、抢险救灾等任务。另外,上述指标还需要根据研究目的进行分解,例如卫生人员分解为医生、护士、其它技术人员等,设备分解为床位、大型设备等。因此,建立包含 100% 全要素的 DEA 模型是不可能的,在实际应用中根据研究目的包含主要的投入产出即可。除模型 1 的投入指标仅包含卫生人员一个要素之外,其它 7 个模型均包含了医院生产的主

要投入和产出要素。

DEA 模型中,投入产出指标数量越多,分析结果的区分度越低。投入或产出指标之间的相关度越低,指标数量对区分度的影响越大。模型 6—8 的产出指标完全相同,区别在于投入指标的卫生人力部分不同。在卫生人力指标中,卫生人员数和其它 3 个指标的相关系数均超过了 0.99,因此模型 6—8 的分析结果差别很小。

对模型区分度的要求没有统一的标准,不同分析目的对效率区分度的要求是不同的,如果分析目的是发现样本中少数效率最低的单元,则在区分度较低的情况下即可满足分析目的;如果目的是发现样本中效率最优的少数单元,则要求更高的区分度。在一般情况下,被评价为有效的单元数量不宜超过总数的 1/3。根据这一标准,模型 1—3 是合适的。但模型 1 的投入仅包括了卫生人力,不符合 DEA 模型应包含主要投入和产出指标的规则。综合判断,以省为单位评价医院效率时,在 8 个模型中,模型 2 和模型 3 的指标选择是最合理的。

DEA 模型指标选择的一个特点是,指标之间的共线性(高度相关)的存在不会对模型的稳定性和可靠性产生影响。^[1]如果在投入(产出)指标中增加一个与已有投入(产出)指标完全线性相关的新指标,分析结果保持不变。因此,在选择投入或产出指标时无需考虑共线性的问题。

3.2 对整体效率和效率前沿的判断

本文用医院效率的实际数据证明,投入产出指标的数量严重影响 DEA 模型的分析结果。在效率区分度不高的情况下(例如模型 6—8 和文献^[18]中的结果),对整体效率和效率前沿的判断是不可靠的。特别是对于判定为有效的决策单元,即效率前沿,由于存在默认有效的问题,需要结合超效率值、被参比次数和总权重等信息进行综合的判断。在数据支持的情况下,应该通过比较多种投入产出指标的组合以检验 DEA 分析结论的可靠性。

3.3 专题地图在区域分析中的作用

在分析某个指标的区域分布特征时,最常用的方式是将全国 31 个省份划分为东、中、西部,这是主要基于经济发展水平的区域划分方式。卫生领域的许多研究指标会受到社会、经济、地理、文化等因素

的共同影响,东、中、西部的区域划分方式未必与研究指标的真实分布状况一致。在没有专门的卫生指标区域划分方法的情况下,应首先采用描述性的方式探索指标的区域分布特征。

专题地图是一种基于地理信息系统的以地图形式直观展示分析指标地理分布特征的方法,它可以为进一步探索效率区域分布特征的内在因素和进行实证研究提供基础。通过专题地图可以明显地发现,我国医院效率的地区分布并没有呈现东、中、西的特征,而是南、中、北的特征。这说明既往研究中在分析效率区域分布特征时先入为主地采用东、中、西的划分方法是不合适的,而是应该先观察实际分布特征,再探索特征背后的相关因素。

参 考 文 献

[1] Jacobs R, Smith P C, Street A. Measuring efficiency in health care: analytic techniques and health policy [M]. Cambridge University Press, 2006.

[2] Hollingsworth B, Peacock S. Efficiency measurement in health and health care [M]. Routledge, 2008.

[3] Charnes A, Cooper W W, Rhodes E. Measuring the efficiency of decision making units [J]. European Journal of Operational Research, 1978, 2(4): 429-444.

[4] Emrouznejad A, Parker B R, Tavares G. Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA [J]. Socio-Economic Planning Sciences, 2008, 42(3): 151-157.

[5] Cook W D, Seiford L M. Data envelopment analysis (DEA) -Thirty years on [J]. European Journal of Operational Research, 2009, 192(1): 1-17.

[6] Sherman H D. Measurement of hospital technical efficiency: A comparative evaluation of data envelopment analysis and other efficiency measurement techniques for measuring and locating inefficiency in health care organizations [D]. United States Massachusetts: Harvard University, 1981.

[7] Hollingsworth B. Non-parametric and parametric applications measuring efficiency in health care [J]. Health Care Management Science, 2003, 6(4): 203-218.

[8] Ozcan Y A. Health care benchmarking and performance evaluation: An assessment using data envelopment analysis (DEA) [M]. Springer, 2007.

[9] Hollingsworth B. The Measurement of Efficiency and Productivity of Health Care Delivery [J]. Health Economics, 2008, 17(10): 1107-1128.

[10] 庄宁. 数据包络分析在国外医院效率评价中的应用 [J]. 国外医学(医院管理分册), 2001(3): 101-105.

[11] 史健, 魏权龄. DEA 方法在卫生经济学中的应用 [J]. 数学的实践与认识, 2004, 34(4): 59-66.

[12] 钟若冰, 张靖, 廖菁. 医院效率评价方法的研究 [J]. 实用医院临床杂志, 2010, 7(1): 127-130.

[13] Cooper W W, Seiford L M, Tone K. Introduction to data envelopment analysis and its uses: with DEA-solver software and references [M]. Springer, 2006.

[14] Ruggiero J. Impact assessment of input omission on DEA [J]. International Journal of Information Technology and Decision Making, 2005, 4(3): 359-368.

[15] 张纯洪, 刘海英. 我国区域卫生经济系统的投入产出技术效率测度研究 [J]. 中国卫生经济, 2009, 28(7): 11-13.

[16] 罗良清, 胡美玲. 中国各地区医疗卫生服务的生产效率分析 [J]. 统计与信息论坛, 2008, 23(2): 47-51.

[17] 刘军, 钱力. 我国区域经济运行效率研究——以医疗卫生系统为例 [J]. 经济问题, 2011(6): 114-118.

[18] 张瑞华, 刘莉, 李维华, 等. 基于数据包络分析的我国 31 个省市医疗卫生服务效率评价 [J]. 中国卫生经济, 2011(2): 69-72.

[19] 张宁, 胡鞍钢, 郑京海. 应用 DEA 方法评测中国各地区健康生产效率 [J]. 经济研究, 2006(7): 92-105.

[20] 杨文芳. 我国卫生资源的省际分布与利用效率研究 [J]. 中国卫生资源, 2010(1): 8-10.

[21] Emrouznejad A, Amin G R. DEA models for ratio data: Convexity consideration [J]. Applied Mathematical Modelling, 2009, 33(1): 486-498.

[22] Gravelle H, Wildman J, Sutton M. Income, income inequality and health: What can we learn from aggregate data? [J]. Social Science & Medicine, 2001, 54(4): 577-589.

[23] Hollingsworth B, Wildman J. The efficiency of health production; re-estimating the WHO panel data using parametric and non-parametric approaches to provide additional information [J]. Health Economics, 2003, 12(6): 493-504.

[收稿日期:2012-01-15 修回日期:2012-02-20]

(编辑 薛云)