

# 基于德尔菲法的公立医院绩效评价体系筛选研究

葛爱晨<sup>1</sup>, 陈家应<sup>1</sup>, 季国忠<sup>2</sup>, 姜荣勤<sup>1</sup>, 李静娴<sup>1</sup>

(1.南京医科大学医政学院, 江苏 南京 211166; 2.南京医科大学第二临床医学院, 江苏 南京 210011)

**摘要:**目的:采用德尔菲法编制基于患者体验的公立医院绩效评价量表。方法:在参考国内外基于患者体验的医院绩效评估理论、方法、效果量表的基础上,结合对江苏省5城市15家公立医院(其中5家城市公立医院,10家县级公立医院)共180名患者(门诊、急诊和住院患者)的访谈情况,总结归纳出基于患者体验的绩效指标库,并根据统计出的指标和指标出现的频数(即词频)设计出本次指标体系的初始量表,通过两轮专家咨询及一轮专家座谈会量表进行进一步的筛选、优化,形成较为合理和有效的量表。结果:通过三轮德尔菲法对量表条目进行筛选,初步形成由“医患关系”、“服务管理”、“质量效率”、“医疗费用”及“总体评价”5个维度,16个方面组成的评价指标体系。结论:该指标体系对基于患者体验的公立医院绩效评价具有一定的参考价值。

**关键词:**德尔菲法;指标体系;公立医院;绩效评价

中图分类号:R197.3

文献标志码:A

文章编号:1671-0479(2017)03-227-005

doi:10.7655/NYDXBSS20170313

随着我国医疗卫生体制改革的不断深化,医疗服务市场化日趋显现,医疗卫生界逐渐接受了患者作为消费者的概念<sup>[1]</sup>。近年来,医院绩效评估研究的重点已从医院的服务能力、服务过程转向了以患者为中心,医疗行业“以人为本”、“以患者为中心”理念逐渐深入人心。如何科学评价医疗卫生服务、提高医疗服务质量、获得群众的认可,成为当今卫生相关部门关心的主要话题。以技术质量、技术效率、服务行为等服务提供的评价体系,更多地关注医疗服务的技术环节、服务过程、服务态度等,而作为服务接受者的患者感受,除了满意度外,缺乏其他有关的测量工具<sup>[2]</sup>。开发基于患者就医体验的评价量表,对测量患者对医疗卫生改革的获得感,以反映医药卫生改革对公立医院服务绩效的影响,很有必要。

为了能有效评估公立医院绩效,本课题组拟研制基于患者体验的公立医院绩效评价量表。文章在参考国内外有关患者报告体验(patient-reported experience measures, PREMs)的理论、方法、效果的基

础上<sup>[3-5]</sup>,结合对患者的访谈情况,总结归纳出基于患者体验的绩效指标库,并筛选相关指标制定专家咨询表,运用改良德尔菲专家咨询法<sup>[6-7]</sup>,开展公立医院基于患者体验的服务绩效评价指标体系。

## 一、资料与方法

### (一)建立指标库

由6人组成课题小组,查阅和复习国内外相关文献,通过偶遇法对江苏省5城市15家公立医院(其中5家城市公立医院,10家县级公立医院)180名患者(门诊、急诊和住院患者)进行访谈。访谈的目的在于通过对有就医体验或认知的患者、医务人员以及一般人群的面对面交谈,获得目前患者对医疗服务需求和期望的第一手资料。经过统计分析初步得出与研究相关的包含7个维度、共140个指标的备选指标库,并在此基础上初步确定5个维度17个二级指标的指标体系。

### (二)德尔菲法

本研究为了提高德尔菲法的效率,对其进行了

**基金项目:**国家自然科学基金项目“基于患者体验的公立医院改革绩效评估研究”(71473129)

**收稿日期:**2017-03-07

**作者简介:**葛爱晨(1991—)男,江苏盐城人,硕士研究生在读;季国忠(1964—),男,江苏南通人,教授,研究方向为消化内科和医院管理,通信作者。

改良,以两轮专家咨询、一轮专家座谈会的形式开展调查。国内完全基于患方的体验开展类似的研究几乎空白,也尚未有完整的评价体系,所以很难通过2~3轮的咨询达到高度的专家共识。因此,开展一轮专家座谈会,并且参加座谈会的专家也基本是前2轮专家咨询的成员,这样的咨询对于快速形成量表有较大的帮助。根据本研究的主要目的,遵循权威性和代表性相结合、多学科交叉的原则,根据专家的知识领域、职称、学历及工作年限等条件遴选专家,所选专家必须为卫生行政管理者、医院行政管理或卫生事业管理研究方面的专家,从事与本课题相关的工作10年以上,在相关领域具有一定的权威性和代表性。有关研究表明<sup>[8]</sup>,德尔菲法的专家人数控制在15~50人为宜,经过小组讨论后确定选取28位符合以上条件的专家进行咨询。

### (三)编制专家咨询表

课题小组根据对备选指标库出现的各指标频数(词频)的统计,通过讨论、分析研究主题,初步确定“医患关系”、“服务管理”、“质量效率”、“医疗费用”、“满意度”这5个维度及其所涵盖的二级指标,制定第一轮专家咨询表。由专家对每个一级和二级指标的重要性、可行性及判断依据(理论分析、实践经验、参考文献)作出评价。其中重要性和可行性按1~9分进行打分<sup>[9]</sup>,1分为最不重要或最不可行,9分为最重要或最可行;判断依据按1~5分打分,1分为依据强度最低,5分为依据强度最高。专家也可根据所在领域的经验修改指标或者补充其他指标。第二轮专家咨询表在根据第一轮专家意见进行修改完善的基础上,附上各专家的意见以及课题组的讨论结果。

### (四)质量控制

在咨询前开始联系专家,确保能够参加两轮以上的专家咨询;咨询开始后,在回收截止日期前,通过邮件、电话等途径进行提醒,提高应答率。

### (五)统计学方法

采用SPSS 20.0软件对数据进行统计分析。结果中给出专家积极系数(问卷回收率)、权威系数(专家对指标的熟悉程度)、反映专家意见集中程度的指标(均数、标准差)及反映专家意见协调程度的指标(变异系数)<sup>[10]</sup>。

## 二、结果

本研究共邀请28位专家进行两轮咨询,专家工作领域分为三个层面:卫生行政管理者、医院行政管理和卫生事业管理研究。第一轮咨询发放问卷

28份,回收27份,有效问卷27份,问卷回收率为96.43%。第二轮咨询发放问卷28份,回收26份,有效问卷26份,问卷回收率为92.86%。在此基础上邀请了10位专家以专家座谈会的形式,对指标进行调整和修改,并对最终确立的指标进行权重赋值。

### (一)专家的基本情况

结果显示,从咨询专家学历来看,最多的是博士(第一轮70.37%,第二轮76.92%),其次是硕士(第一轮18.52%,第二轮19.23%),专家团队整体以高学历为主;从职称来看,90%以上专家具有副高以上职称(第一轮92.59%,第二轮96.15%);从工作领域来看,专家以从事卫生事业管理为主(第一轮77.78%,第二轮76.92%);从工作年限来看,所有专家在课题相关领域工作均超过12年。第一轮专家咨询反馈的27份咨询表和第二轮专家咨询反馈的26份咨询表分析得到的专家基本情况见表1。

表1 专家基本情况

项 目	第一轮		第二轮	
	人数	构成比(%)	人数	构成比(%)
学历				
本科	2	7.41	1	3.85
硕士	5	18.52	5	19.23
博士	19	70.37	19	76.92
其他	1	3.70	1	3.85
职称				
中级	1	3.70	1	3.85
副高	3	11.11	3	11.54
正高	22	81.48	22	84.61
其他	1	3.70	0	0.00
从事的主要工作				
卫生行政管理	2	7.41	1	3.85
医院管理	4	14.81	5	19.23
卫生事业管理	21	77.78	20	76.92
从事专业领域年限(年)				
12~20	8	29.63	8	30.77
>20~30	10	37.04	9	34.62
>30~38	9	33.33	9	34.61

### (二)专家咨询的可靠性

#### 1.积极系数

专家积极系数体现专家对该研究的关心程度,可通过专家咨询问卷的有效回收率来表示。本研究中第一轮专家咨询回收率为96.43%,第二轮专家咨询回收率为92.86%。在第一轮中,51.85%的专家(14/27)对指标体系提出了改进建议,在第二轮中,46.15%(12/26)的专家对指标体系提出了改进建议。结果表明,整体上专家对本研究比较支持和关心。

## 2.协调系数

专家意见的协调程度用 Kendall 协调系数 W 表示,并通过检验推断各专家意见是否具有-致性。协调系数 W 在 0~1 之间,越接近于 1 表示专家意见协调程度越高。结果显示,第一轮及第二轮专家对指标重要性和可行性的检验结果均有统计学意义 ( $P<0.001$ ),表明专家对指标的评价具有一致性(表 2)。

表 2 专家意见协调系数

指标	第一轮		第二轮	
	重要性	可行性	重要性	可行性
W	0.267	0.189	0.181	0.162
$\chi^2$	202.010	143.250	98.714	88.511
P	<0.001	<0.001	<0.001	<0.001

## 3.专家权威程度

专家的权威程度可通过专家的判断依据来间接体现。专家根据自己判断依据的情况,对每个指标的理论分析、实践经验和参考文献按 1~5 分进行打分,1 分为最低,5 分为最高,分值越高表示依据强度越强。结果表明,卫生行政管理者和医院管理者的判断依据主要是实践经验,其次是理论分析,这是由于他们大部分来自各医疗机构,工作领域更侧重于管理实践,具有更丰富的实践经验。而卫生管理研究专家的判断依据主要是理论分析,其次是

实践经验,这是由于其大多来自高校等学术机构,工作领域更侧重于理论分析。对判断依据的分析结果与专家从事的工作性质相符合(表 3)。

表 3 专家权威程度 (分)

专家研究领域	第一轮			第二轮		
	理论分析	实践经验	参考文献	理论分析	实践经验	参考文献
卫生行政管理者	3.22	4.59	2.40	2.70	4.61	2.61
医院管理者	3.56	4.14	2.15	3.34	3.59	3.11
卫生管理研究专家	4.19	3.67	2.81	4.08	3.93	3.19

### (三)专家意见的集中程度和协调程度

从均数、标准差、变异系数这三个指标来考察指标的重要性和可行性。其中,均数越大,表明指标的重要性越高、可行性越好;标准差越小,表明专家意见越集中;变异系数越小,表明专家对该指标的协调程度越高。

专家咨询第一轮一级指标的重要性赋值均数最小值为 7.41 分,最大值为 8.52 分;可行性赋值最小值为 7.22 分,最大值为 7.85 分;变异系数均小于 20%;第二轮一级指标的重要性赋值均数最小值为 7.63 分,最大值为 8.40 分;可行性赋值最小值为 7.02 分,最大值为 7.87 分;变异系数均小于 20%(表 4)。

表 4 一级指标专家咨询结果(第一轮、第二轮)

指标	第一轮				第二轮			
	重要性赋值(分)	变异系数(%)	可行性赋值(分)	变异系数(%)	重要性赋值(分)	变异系数(%)	可行性赋值(分)	变异系数(%)
医疗费用	8.19±0.83	10.19	7.85±1.23	15.68	8.06±0.88	10.87	7.87±1.21	15.42
医患关系	7.96±1.02	12.79	7.22±1.37	18.94	7.83±1.10	14.11	7.02±1.06	15.14
质量与效率	8.52±0.89	10.48	7.85±1.10	14.00	8.40±1.06	12.60	7.62±1.13	14.89
服务管理	7.41±1.36	17.67	7.33±1.27	17.33	7.63±1.20	15.67	7.33±1.19	16.26
满意度	8.26±0.86	10.40	7.63±1.28	16.72	7.93±1.20	15.12	7.73±0.96	12.44

专家咨询第一轮二级指标的重要性赋值均数最小值为 6.19 分,最大值为 8.74 分;可行性赋值最小值为 6.77 分,最大值为 8.60 分;3 个指标的变异系数大于 25%;第二轮二级指标的重要性赋值均数最小值为 6.63 分,最大值为 7.83 分;可行性赋值最小值为 6.00 分,最大值为 8.33,5 个指标的变异系数大于 25%(表 5)。

参考相关文献后将指标的删除标准定为可行性小于 7 分,标准差大于 1 分,变异系数大于 25%。当指标达到删除标准时,将其纳入考虑删除的范围,进一步结合专家意见和各指标的内容和特点,经课题组讨论后对指标体系进行相应的调整。经两

轮专家咨询后,最终得到一级指标 5 个,二级指标 15 个。

有专家认为门诊患者很少在医院食堂吃饭,因此“餐食服务”的适用性不强,讨论后删除;有专家认为“患者信任”、“流程效率”表达欠妥,课题组考虑到患者信任是单方面的,但在就医过程中其实是一种医患双方互相信任的关系,将“患者信任”改名为“医患互信”,将“流程效率”改名为“服务流程”。

对于部分指标的布局作修改。有专家认为“治疗效果”需放在第一位,以突显治疗效果的重要性,结合患者调查的词频分析,课题组讨论将“治疗效果”调至“质量与效率”栏目下的第一个二级指标;

表5 二级指标专家咨询结果(第一轮、第二轮)

一级指标	二级指标	第一轮				第二轮			
		重要性 赋值(分)	变异系 数(%)	可行性 赋值(分)	变异系 数(%)	重要性 赋值(分)	变异系 数(%)	可行性 赋值(分)	变异系 数(%)
医疗费用	就医总费用	8.44±0.80	9.48	8.13±0.82	10.07	7.79±1.30	16.66	8.07±1.24	15.33
	医生处方行为	7.89±2.06	26.15	7.50±1.03	13.73	6.73±1.43	21.24	8.33±0.78	9.41
医患关系	服务态度	7.19±1.69	23.49	8.13±1.00	12.24	7.65±1.16	15.21	6.00±1.94	32.36
	医患沟通	8.27±0.92	11.11	8.08±1.41	17.47	7.32±1.31	17.95	7.85±1.29	16.40
	关心患者	8.04±1.37	17.07	7.82±1.30	16.60	7.06±1.06	15.07	7.48±1.45	19.39
	患者需求响应	8.26±0.76	9.25	7.75±1.18	15.19	7.12±1.07	15.05	7.48±1.55	20.76
质量与效率	患者信任	7.59±1.28	16.84	7.48±1.62	21.59	6.63±1.45	21.90	7.04±1.76	25.08
	技术水平	7.22±1.34	18.55	8.60±0.87	10.07	7.24±1.05	14.53	6.44±1.65	25.58
	流程效率	8.74±0.53	6.01	7.81±1.23	15.72	7.60±1.00	13.16	8.00±0.78	9.81
服务管理	治疗效果	8.26±0.90	10.93	8.52±0.70	8.21	7.60±1.33	17.47	7.89±0.93	11.84
	基础设施	7.96±1.09	13.70	7.80±1.17	14.95	7.73±1.25	16.18	7.41±1.42	19.19
	就医环境	6.63±1.64	24.80	8.02±1.15	14.38	7.83±0.73	9.38	6.48±1.65	25.45
	餐食服务	6.93±1.62	23.32	6.77±1.31	19.29	7.42±1.36	18.34	6.70±1.32	19.76
	人性化服务	6.73±2.24	33.24	7.50±1.24	16.55	6.92±1.38	19.98	6.69±1.98	29.51
满意度	安全管理	6.19±1.94	31.40	7.44±1.64	21.98	7.20±1.47	20.44	7.07±1.47	20.72
	服务评价	7.78±1.28	16.47	8.17±1.05	12.85	7.77±1.02	13.14	7.78±1.09	13.96
	推荐医院	7.56±1.48	19.54	7.18±1.31	18.30	7.18±1.55	21.55	7.78±1.28	16.47

“基础设施”和“人性化服务”存在交叉,所以将“基础设施”并入“人性化服务”;有专家认为医患关系下的各个二级指标间存在交叉、重叠,经课题组结合WHO的卫生系统反应性讨论,将医患关系中的内容重新归并、整合,将二级指标调整为“患者需求响应”、“尊重患者”、“医患沟通”和“隐私保护”。

(四)指标权重赋值

在前两轮专家的修改意见基础上,第三轮专家咨询以座谈会的形式展开。专家经讨论、分析、权衡后,增加1个二级指标,并对各二级指标进行最终定义及命名,初步确立基于患者体验的公立医院绩效评价指标体系(5个一级指标,16个二级指标),并对相应的指标进行权重赋值,命名及赋值情况见表6。

三、讨论

(一)确定了评价体系的框架和主要指标

研究通过三轮专家咨询,确立了5个一级指标、16个二级指标的评价指标体系。包括医患关系中的疾病认知、需求响应、尊重患者、患者沟通、隐私保护;服务管理中的就医环境、安全管理、服务流程、多元服务;质量效率中的治疗效果、技术水平、服务效率;医疗费用中的就医费用、诊疗行为;总体评价中的服务评价以及推荐医院。这一体系经过3轮咨询得出,具有较高的认可度。

表6 一级、二级指标权重赋值(第三轮)

一级指标	赋值(分)	二级指标	权重
医患关系	20	疾病认知	20
		需求响应	20
		尊重患者	20
		医患沟通	25
		隐私保护	15
服务管理	25	就医环境	20
		安全管理	25
		服务流程	35
		多元服务	20
质量效率	25	治疗效果	45
		技术水平	30
		服务效率	25
医疗费用	15	就医费用	45
		诊疗行为	55
总体评价	15	服务评价	65
		推荐医院	35

(二)德尔菲法的专家选择具有代表性

目前国内缺乏基于患者体验评价公立医院绩效的工具或指标体系,本研究基于德尔菲法,建立一套具有普适性的指标体系<sup>[11]</sup>,对公立医院的绩效进行评价。此次专家咨询,对于咨询专家的选择注重权威性和代表性,包括卫生部门、科研机构、医疗领域等代表,并对专家积极系数、协调系数和权威

程度进行了分析,整体上来说本次的专家咨询法是相对可靠、可信的。

### (三)评价指标体系仍存在争议与不足

通过德尔菲专家咨询法初步得到了评价指标体系,但是在分析中明显存在一些方面的专家意见不统一,有专家在费用指标中发生分歧,有专家建议使用“医疗费用”,也有专家觉得“医疗经济负担”更为恰当,讨论之后认为患者就医过程中直接感知的是费用而不是负担,先确定为“医疗费用”;有专家建议将“患者需求响应”改为“合理需求响应”,但考虑到从患者角度考虑其所有需求都是合理的,最终暂定为“患者需求响应”;有专家认为患者自付的费用才是患者体验最核心的,建议将“就医总费用”改为“患者自付费用”,考虑到部分患者就诊过程中无法获知自付费用明细,暂定“就医总费用”。对于存在的矛盾与争议,可以考虑在接下来的实证性研究,评估这些指标在调查中是否具有可行性和结果是否具有真实性,并在此基础上,进一步修改完善指标体系。

### 参考文献

- [1] 黄银芳. 患者视角下的医院住院服务质量评价和改进研究[D]. 杭州:杭州师范大学,2013
- [2] 郭静竹,朱志峰,赵红梅,等. 北京某综合医院住院患者满意度调查对比研究[J]. 中国医院管理,2009,29(12):56-58
- [3] 王江蓉,黄森,Thomas S,等. 中国部分医院患者体验和满意监测工具应用比较分析 [J]. 北京大学学报(医学版),2011,43(3):397-402
- [4] Labarère J,François P,Bertrand D,et al. Outpatient satisfaction:validation of a French -language questionnaire:data quality and identification of associated factors [J]. Clin Perform Qual Health Care,2000,7(2):63-69
- [5] John S. How valid and reliable are patient satisfaction data:an analysis of 195 studies [J]. International Journal for Quality Health Care,1999,11(4):319-328
- [6] 龙耀,周璇,邓鹏飞,等. 基于改良德尔菲法的健康人生理性衰老量表构建 [J]. 中南大学学报(医学版),2016,41(4):422-426
- [7] Gwee X,Nyunt MS,Kua EH,et al. Reliability and validity of a self -rated analogue scale for global measure of successfulaging[J]. Am J Geriatr Psychiatry,2014,22(8):829-837
- [8] 谈立峰,郝东平,孙樾陵,等. Delphi 法在评价区域性大型活动公共卫生突发事件风险发生的可能性及后果中的应用[J]. 职业与健康,2012,28(10):1160-1164
- [9] Li Q. A novel Likert scale based on fuzzy sets theory[J]. Expert Syst Appl,2013,40(5):1609-1618
- [10] 孙振球. 医学统计学 [M]. 北京:人民卫生出版社,2002:376-378
- [11] 陈英耀,倪明,胡献之,等. 公立医疗机构公益性评价指标筛选——基于德尔菲专家咨询法[J]. 中国卫生政策研究,2012,5(1):6-10