

责编意见： 1、作者信息需要修改。2、讨论部分二级标题要概括出主要观点。3、建议是针对样本医院的，讨论是针对指标体系的，不对应。建议部分增加关于指标体系的内容，实践层面是为了验证指标体系的科学合理性，不要过度发散。

DRG 支付改革背景下基于价值医疗导向的医疗服务绩效评价指标构建及应用——以 X 省三级医院为例

马佳妮^{1,*}, 马亦晨^{1,*}, 杜旦², 韩雪梅¹, 张雨佳¹, 姚璇¹, 万婉¹, 郭卓咏¹, 任育彪¹, 韩雪梅^{1,*}, 刘艳英^{2,*}

1. 兰州大学公共卫生学院 兰州 730000

2. 兰州大学第一医院 兰州 730000

注：“*”代表马佳妮与马亦晨为共同第一作者；“*”代表韩雪梅与刘艳英为共同通讯作者。

【摘要】目的：在 DRG 支付方式改革背景下，构建以价值医疗为导向的医疗服务绩效评价指标体系，以 X 省四所三级医院为例，综合分析不同维度医疗服务绩效水平差异。**方法：**采用德尔菲专家咨询法构建指标体系，通过博弈论法整合专家评分法、熵值法、CRITIC 法进行组合赋权；运用综合指数法、秩和比法（RSR）、逼近理想解排序法（TOPSIS）开展绩效评价，经 Spearman 等级相关分析验证三种方法的一致性。**结果：**经两轮专家咨询，最终形成含 6 个一级指标、22 个二级指标的评价体系；组合赋权后各指标权重差异较小；三种综合评价方法结果显示，四所样本医院在各维度绩效水平存在显著差异。**结论：**以价值医疗为导向构建的评价指标体系更加客观全面；指标权重设置科学合理、引导性强；三种综合评价方法可有效反映医院绩效差异，为 DRG 支付改革背景下的医疗服务绩效评价提供了多维度实证依据与实践参考。

【摘要】目的：在 DRG 支付方式改革背景下，构建以价值医疗为导向的医疗服务绩效评价指标体系，以 X 省四所三级医院为例，综合评价分析样本医院在不同维度的医疗服务绩效水平差异。**方法：**运用德尔菲专家咨询法构建三级医院医疗服务绩效评价指标体系，并运用博弈论法，结合专家评分法、熵值法、CRITIC 法（Criteria Importance Though Intercriteria Correlation）对所得结果进行组合赋权，并运用综合指数法、秩和比法（Rank Sum Ratio, RSR）、逼近理想解排序法（Technique for Order Preference by Similarity to an Ideal Solution, TOPSIS）三种综合评价方法对 DRG 医疗服务绩效进行评价，经 Spearman 等级相关分析法验证三种综合评价方法的一致性。**结果：**经两轮德尔菲专家咨询，构建基于 DRG 的医疗服务绩效评价指标体系，共包括 6 个一级指标，22 个二级指标，运用博弈论法进行组合赋权，结果显示各指标间权重系数差异较小，其中权重最大的指标为费用消耗指数（0.069），权重最小的指标为病例组合指数值（Case Mix Index, CMI）（0.033），三种综合评价方法分析四所医院医疗服务绩效评价结果在各维度上存在差异。**结论：**以价值医疗为导向构建的评价指标体系更加客观全面，评价体系的指标权重设置进一步合理化，科学性、引导性增强，三种评价方法能客观反映医院绩效差异，为 DRG 改革下的医疗绩效评价提供多维度实证依据。

【关键词】 医疗服务绩效评价；博弈论法；价值医疗；DRG

中图分类号：R197.1、R197.3 文献标志码：A

Construction and Application of Value-Based Healthcare Oriented Performance Evaluation Indicators for Medical Services under the DRG Payment Reform: A Case Study of Tertiary Hospitals in Province X

MA Jiani^{1,*}, MA Yichen^{1,*}, HAN Xuemei¹, DU Dan², ZHANG Yujia¹, YAO Xuan¹, WAN Wan¹, GUO Zhuoyong¹, REN Yubiao¹, HAN Xuemei^{1,*}, LIU Yanying^{2,*}

1. School of Public Health, Lanzhou University, Lanzhou 730000, P.R. China

2. The First Hospital of Lanzhou University, Lanzhou 730000, P.R. China

1项目基金：甘肃省哲学社会科学规划项目阶段性研究成果（2023YB056）***作者简介：**马佳妮（2002 年—），女，硕士研究生，主要研究方向为医药卫生方针政策与法律法规研究。E-mail: majin2024@lzu.edu.cn

马亦晨（1999 年—），女，硕士，主要研究方向为医院绩效管理。E-mail: 220220913520@lzu.edu.cn

***通讯作者：**韩雪梅，E-mail: xmhan@163.com；刘艳英，E-mail: haozhang312@126.com；

【 Abstract 】 Objective: Against the backdrop of DRG (Diagnosis-Related Groups) payment reform, to construct a value-based healthcare-oriented performance evaluation indicator system for medical services, and to comprehensively analyze the differences in medical service performance across various dimensions using four tertiary hospitals in Province X as case studies. **Methods:** The Delphi expert consultation method was employed to construct the indicator system. The game theory method was used to combine the weighting results obtained from expert scoring, the entropy weight method, and the CRITIC method. Three comprehensive evaluation methods—the Comprehensive Index Method, the Rank Sum Ratio (RSR) Method, and the Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS)—were applied to evaluate medical service performance. The consistency of the results from these three methods was verified using Spearman rank correlation analysis. **Results:** After two rounds of Delphi expert consultation, a DRG-based medical service performance evaluation indicator system was established, consisting of 6 first-level indicators and 22 second-level indicators. After combined weighting using the game theory method, the differences in weight coefficients among the indicators were relatively small. The results of the three comprehensive evaluation methods showed significant differences in the performance levels of the four sample hospitals across various dimensions. **Conclusions:** The evaluation indicator system constructed under the guidance of value-based healthcare is more objective and comprehensive. The indicator weighting mechanism is scientifically sound and well-targeted. The three evaluation methods effectively reflect differences in hospital performance, providing multidimensional empirical evidence and practical references for medical service performance evaluation under the context of DRG payment reform.

Objective: Against the backdrop of DRG (Diagnosis-Related Groups) payment reform, to construct a value-based healthcare-oriented performance evaluation indicator system for medical services, and to comprehensively evaluate and analyze the differences in medical service performance across various dimensions using four tertiary hospitals in Province X as case studies. **Methods:** The Delphi expert consultation method was employed to construct the performance evaluation indicator system for tertiary hospitals. Game theory was utilized to combine weighting results obtained from expert scoring, entropy weight method, and CRITIC method (Criteria Importance Though Intercriteria Correlation) for combined weighting. Three comprehensive evaluation methods—Comprehensive Index Method, Rank Sum Ratio (RSR) Method, and Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS)—were applied to evaluate DRG medical service performance. The consistency of the results from these three methods was verified using Spearman rank correlation analysis. **Results:** After two rounds of Delphi expert consultation, a DRG-based medical service performance evaluation index system was constructed, including 6 first-level indicators and 22 second-level indicators. The game theory method was used for combined weighting, and the results showed that the difference in weight coefficients among the indicators was small. Among them, the indicator with the highest weight was the cost consumption index (0.069), and the indicator with the lowest weight was the Case Mix Index (CMI) (0.033). The results of the three comprehensive evaluation methods indicated that there were differences in the medical service performance evaluation results of the four hospitals across various dimensions. **Conclusions:** The evaluation index system constructed under the guidance of value-based healthcare is more objective and comprehensive. The indicator weighting mechanism has been further rationalized, strengthening both scientific rigor and guidance effectiveness. The three evaluation methods objectively reflect hospital performance variations, collectively providing multidimensional empirical evidence for medical performance assessment under DRG reform.

【 Keywords 】 Performance Evaluation of Medical Services; Game Theory Method; Value-Based Healthcare; Diagnosis Related Groups

当下，我国医药卫生体制改革已经进入了新发展阶段，需要积极推动医药卫生体制从服务量驱动向价值驱动转变^[1]，而价值医疗理念与公立医院高质量发展、DRG支付方式改革的目标高度契合^[2]，即促使医院不断优化收入结构、提升效率、降低成本^{[3][4]}、提高医疗质量以及患者满意度^[5]。本研究推动价值医疗理念融入医疗服务绩效评价指标的构建，通过多种综合评价方法进行医疗服务绩效测度与对比分析，为DRG改革背景下医疗绩效评价提供多维度、可比较的实证支撑，助力实现医、保、患三方共赢。

1 资料与方法

1.1 数据来源

本研究数据主要来源于X省医疗保障局，所选取的4所样本医院均为该省首批DRG/DIP示范医院，且均已实际实施DRG支付方式改革，具备典型代表性。本研究中数据主要来源于X省医疗保障局。从病案首页调取四所样本医院2024年1月至11月以DRG

方式进行结算的数据，按照筛选标准纳入相关的患者基本信息、就医信息、临床特征及相关费用信息，剔除不在入组范围内的病例、重复病例、关键信息缺失及住院天数及费用存在明显异常的病例，最终在后续分析时保留了 177 768 例高质量的数据，确保研究结果具有准确性及可靠性。其中，A 医院共 43 497 例，B 医院 38 853 例，C 医院 45 580 例，D 医院 49 838 例。

1.2 研究方法

1.2.1 德尔菲专家咨询法

德尔菲专家咨询法(Delphi Technique)是根据相关专家的直接经验，经过反复征询、归纳、修改，最终将专家的意见汇总为预测结果，最终形成一个较为统一且可靠性较高的结论或方案^{[6][7]}，具有匿名性、科学性及可操作性的属性，目前已广泛用于设计量表^[8]。

本研究运用德尔菲专家咨询法构建甘肃省三级医院医疗服务绩效评价指标体系本研究在构建绩效评价指标体系时，将价值医疗理念转化为可以被量化评价的框架，首先通过主题检索梳理相关国内外文献及公立医院高质量发展评价指标（试行）操作手册，参考三级公立医院绩效考核指标，确定初始指标池共包括 6 个一级指标、33 个二级指标。之后，进行两轮专家咨询，邀请相关领域工作 5 年以上，且具备扎实的理论基础和专业工作经验的来自甘肃省医疗保障局、北京中医药大学管理学院、北京广安门医院、浙江大学医学院附属邵逸夫医院、四川大学华西医院等 10 个单位的 14 位专家发放问卷参与本次专家函询工作，逐步论证并筛选绩效评价指标，兼顾专家的代表性，坚持理论与实践相结合，确保此次绩效评价指标体系的构建具有科学性。

1.2.2 确定指标权重的方法

本研究在确定权重时综合应用多种方法，具体包括：（1）基于专家评分法确定主观权重，依托第二轮专家咨询中 12 名专家对各指标的相对重要性的评分，计算指标体系中各指标的主观权重系数^[9]；（2）基于熵值法确定客观权重，通过指标值的离散程度来计算信息熵，从而量化各评价指标在体系中的影响力^[10]；（3）CRITIC 法(Criteria Importance Though Intercriteria Correlation)确定客观权重，通过标准差计算指标数据的离散程度量化指标的“辨别力”；通过计算指标间的相关系数来判断指标的“冲突性”^[11]；（4）基于博弈论法确定组合权重，将单一赋权结果视为博弈参与方，通过矩阵运算建立博弈方程组，继而求解得到能够最大限度协调主客观信息矛盾的组合权重^[12]，将上述三种权重结果组合赋权，从而确定绩效评价指标体系的最终权重。

1.2.3 综合评价方法

（1）综合指数法

综合指数法是运用统计学方法，将评价对象的多个性质、计量单位不同的指标科学全

面地转化为一个综合指数。依据该方法能使 DRG 指标间的可比性更好，可以评价绩效水平的整体优劣^{[13][14]}。计算得出的综合指标值越大，表明医疗服务绩效水平越高。

(2) 秩和比法(Rank Sum Ratio, RSR)

其原理在于通过秩将一个 n 行 m 列的矩阵转换为无量纲统计量 RSR，得到的 RSR 值介于 $0\sim 1$ 之间，该值越大说明评价对象综合评价越“优”，反之则表明评价对象越“劣”。之后，运用参数统计分析的概念和方法，继而运用回归模型得到 RSR 值的分布，通过 RSR 拟合值计算实现评价对象的精准排序或等级划分^[15]。

(3) TOPSIS 法

其核心理念就是从多个方案中选出各因素的最优值与最劣值，其中正理想解 (D+) 是由最优值组成的解，负理想解 (D-) 是由最劣值组成的解，将各方案中的正、负理想解之间的欧氏距离 (Ci) 进行计算，再通过欧氏距离确定各方案的优劣，进行排序^[16]。

2.结果

2.1 问卷及专家情况

2.1.1 问卷信效度检验

本研究采用 SPSS 25.0 软件进行信度检验，分析专家函询表的评分结果，一级指标与二级指标总体的克隆巴赫 α 系数分别为 0.710、0.781，大于 0.7 这一界限值^[17]，表明符合问卷信度条件；本研究兼顾理论与实践满足德尔菲专家咨询法的要求，通过严密地设计、修改问卷，促使问卷具有良好的效度。

2.1.2 专家积极程度

最终发出咨询问卷表 14 份，有效回收问卷数 12 份，第一轮与第二轮专家回复率分别达 85.714%、100%，均高于 70%则认为专家积极系数较高^[18]，且两轮有效回收率均达 100%。

2.1.3 专家权威程度

本研究中第一轮专家咨询中 $Cr=0.808$ ，第二轮专家咨询中 $Cr=0.819$ ，该值均大于 0.7，表明本研究有较高的专家权威程度^[19]。

2.1.4 专家意见协调系数

专家意见协调系数一般是判断专家对指标评分的一致性水平，本研究专家咨询结果协调系数见表 1，经检验后差异均具有统计学意义 ($P<0.05$)，表明专家意见的协调性较好，指标分布合理，结果可取^[20]。

表 1 重要性得分的协调系数与显著性检验结果

咨询次数	指标	<i>N</i>	<i>Df</i>	<i>W</i>	κ^2	<i>P</i>
第一轮	一级指标	12	5	0.200	12.018	0.035
	二级指标	12	32	0.239	68.732	0.001
第二轮	一级指标	12	5	0.324	19.426	0.002
	二级指标	12	24	0.242	69.807	0.000

2.2 指标确定

以重要性为依据，根据专家评分结果计算出算术均数、变异系数及满分比，利用界值法筛选绩效评价指标。经过两轮函询专家意见的整理修正后，最终完善和确立了基于 DRG 的医疗服务绩效评价指标体系，共 6 个一级指标、22 个二级指标，见表 2。此外有专家指出“好转率”和“治愈率”因数据计算困难且未纳入省内医保系统，故未纳入实证分析；此外，医院学科骨干及学科带头人数量因统计口径不统一导致数据差异显著，为避免评价偏倚，亦未采用相关数据，仅保留其他关键指标进行实证。

表 2 DRG 医疗服务绩效评价指标体系

一级指标	二级指标	评价内容	指标类型
A 医疗服务能力	A1 收治病例覆盖 DRG 组数	治疗病例所覆盖疾病类型范围	高优
	A2 DRG 总权重	反映医疗服务总量	高优
	A3 病例组合指数值 (CMI 值)	医院诊疗能力和技术难度水平	高优
	A4 疑难病例占比 (RW \geq 2)	反映疑难危重疾病	高优
	A5 三四级手术例数占比	出院患者施行三四级手术台次数占同期出院患者手术台次数比例	高优
B 医疗服务效率	B1 平均住院费用	一定时期出院患者平均住院费用	低优
	B2 次均药品费用	一定时期出院患者平均药品费用	低优
	B3 次均耗材费用	一定时期出院患者平均耗材费用	低优
	B4 平均住院日	一定时期出院患者平均住院时间	低优
	B5 费用消耗指数	治疗同类疾病所花费的费用	低优
	B6 时间消耗指数	治疗同类疾病所花费的时间	低优
C 医疗质量安全	C1 治愈率	反映治愈的患者比例	低优
	C2 好转率	反映病情好转的患者比例	低优
	C3 死亡率	采用标准化调整后的死亡率	低优
	C4 中低风险组死亡率	中低风险疾病发生死亡的概率	低优
	C5 31 天再入院率	31 天因相同/相关疾病再次入院	低优
D 医保结算清单质量	D1 入组病例数	反映 DRG 入组的病例数量	高优
	D2 DRG 入组率	反映 DRG 入组的比例	高优
E 患者体	E1 出院患者满意度	出院患者对医疗服务的期望与	高优

验		其实际感知的一致性程度	
F 持续发展	E2 个人自付比例	患者个人承担费用比例	低优
	F1 学科骨干（培养及引进）数量	通过内部培养或外部引进等方式拥有的学科骨干人才数量	高优
	F2 学科带头人（培养及引进）数量	通过内部培养或外部引进等方式拥有的学科带头人数量	高优

2.3 博弈论法权重结果

专家评分法、熵值法和 CRITIC 法的归一化系数分别为 0.416、0.205 和 0.379，博弈论组合结果显示各指标间权重系数差异较小，其中权重最大的指标为费用消耗指数（0.069），最小为病例组合指数值（CMI 值）（0.033）。见表 3。

表 3 组合权重结果

一级指标	权重	二级指标	专家评分法	熵值法	CRITIC 法	博弈论法
A 医疗服务能力	0.181	A1	0.046	0.025	0.055	0.045
		A2	0.048	0.011	0.039	0.037
		A3	0.048	0.005	0.030	0.033
		A4	0.048	0.011	0.037	0.036
		A5	0.044	0.042	0.053	0.047
B 医疗服务效率	0.162	B1	0.046	0.067	0.026	0.043
		B2	0.043	0.033	0.045	0.042
		B3	0.043	0.033	0.044	0.041
		B4	0.045	0.053	0.044	0.046
		B5	0.048	0.151	0.048	0.069
		B6	0.048	0.151	0.026	0.061
C 医疗质量安全	0.140	C1	0.048	0.011	0.044	0.039
		C2	0.045	0.008	0.071	0.047
		C3	0.043	0.045	0.054	0.048
		C4	0.047	0.084	0.041	0.052
		C5	0.046	0.067	0.052	0.053
D 医保结算清单质量	0.185	D1	0.045	0.033	0.048	0.044
		D2	0.045	0.033	0.048	0.044
E 患者体验	0.161	E1	0.045	0.033	0.045	0.042
		E2	0.044	0.033	0.049	0.043
F 持续发展	0.171	F1	0.043	0.067	0.043	0.048
		F2	0.043	0.009	0.060	0.042

2.4 医疗服务绩效评价结果

本研究在进行综合评价前，先将数据进行高优化、低优化处理，而后进行综合评价。

2.4.1 基于综合指数法的医疗服务绩效评价结果

基于综合指数法的分析结果显示，四所样本医院在医疗服务能力、效率、质量安全、医保结算清单质量及患者体验方面均存在差异，就样本医院综合绩效水平情况而言其综合指数排名依次为：D 医院（6.322）、A 医院（5.760）、B 医院（5.462）、C 医院（5.172）。见表 4、表 5。

表 4 四所样本医院各维度绩效评价结果

医院	综合指数法		RSR 法				TOPSIS 法			排序结果	
	综合指数	排名	RSR 排名	Probit	RSR 拟合值	分档等级	正理想解距离	负理想解距离	相对接近度		
医疗服务能力	A	0.862	3	3	5.000	0.535	2	0.432	0.727	0.627	3
	B	2.111	1	2	5.674	0.693	2	0.434	0.782	0.643	2
	C	1.462	2	1	6.534	0.895	3	0.272	0.795	0.745	1
	D	0.230	4	4	4.326	0.377	2	1.000	0.000	0.000	4
医疗服务效率	A	0.726	3	4	4.326	0.457	2	0.881	0.248	0.220	4
	B	0.714	4	3	5.000	0.558	2	0.609	0.535	0.467	3
	C	1.039	2	1	6.534	0.787	3	0.435	0.827	0.656	1
	D	2.294	1	2	5.674	0.659	2	0.606	0.759	0.556	2
医疗质量安全	A	2.131	1	1	6.534	0.837	3	0.174	0.914	0.840	1
	B	0.662	4	4	4.326	0.430	2	0.756	0.484	0.391	4
	C	0.736	3	3	5.000	0.554	2	0.627	0.597	0.488	3
	D	1.730	2	2	5.674	0.679	2	0.586	0.805	0.579	2
医保结算清单质量	A	0.987	3	1	6.534	0.750	3	0.403	0.770	0.656	2
	B	0.879	4	2	4.681	0.500	2	0.709	0.650	0.478	3
	C	1.009	2	2	4.681	0.500	2	0.770	0.403	0.344	4
	D	1.125	1	1	6.534	0.750	3	0.148	0.901	0.859	1
患者体验	A	1.054	2	2	5.674	0.715	2	0.153	0.871	0.850	2
	B	1.095	1	1	6.534	0.982	3	0.000	1.000	1.000	1
	C	0.926	4	4	4.326	0.297	2	0.713	0.646	0.475	3
	D	0.943	3	3	5.000	0.506	2	0.891	0.164	0.155	4

表 5 四所样本医院医疗服务综合绩效评价结果

综合指数法	RSR 法	TOPSIS 法
-------	-------	----------

医院	综合 指数	排名	RSR 值	排名	\sum	R	ρ	P(%)	Probit 值	正理想 解距离	负理想 解距离	相对接 近度	排序结 果
A	5.760	2	0.625	3	1	2	2	50.0	5.000	0.595	0.667	0.529	3
B	5.462	3	0.640	2	1	3	3	75.0	5.674	0.583	0.665	0.533	2
C	5.172	4	0.668	1	1	4	4	93.8	6.534	0.524	0.727	0.581	1
D	6.322	1	0.552	4	1	1	1	25.0	4.326	0.723	0.647	0.472	4

2.4.2 基于 RSR 法的医疗服务绩效评价结果

医疗服务绩效评价指标的 RSR 值分布见表 4、表 5。本研究采用博弈论法组合赋权确定的指标权重，并进一步编制确定秩和比分布，以概率单位 Probit 为自变量、秩和比值 RSR 为因变量，建立线性回归模型，并通过 RSR 拟合值及 RSR 临界（拟合值）进行区间比较，进而得到分档等级水平，本研究将医疗服务能力情况分为 3 档，其中分档等级 Level 数字越大，则表示等级水平越高，即效应越好。如在医疗服务能力维度的线性回归模型为 $RSR = -0.638 + 0.235 * Probit$ ，结果显示医疗服务能力水平最高的是 C 医院，最低的是 D 医院，处于第 3 档的 C 医院医疗服务能力最好，而第 2 档的 A、B、D 这 3 所医院处于中等水平。

基于 RSR 法的综合绩效结果显示：四所样本医院的 RSR 为 0.552~0.668，四所医院存在差异，医疗服务绩效水平最高为 C 医院，最低为 D 医院，处于第 3 档的 C 医院医疗服务效率水平最好，而处于第 2 档的 A、B、C 这 3 所医院的医疗服务效率水平中等。

2.4.3 基于 TOPSIS 法的医疗服务绩效评价结果

本研究依赖于博弈论组合赋权计算出的指标权重，使用 TOPSIS 法得到四所样本医院的正、负理想解距离及相对接近度，并依据结果进行排序，其中，相对接近度越接近 1，说明该医院医疗服务绩效水平则越高^[21]，如本研究中医疗服务能力维度的相对接近度为 0.000~0.745，排名依次为 C 医院、B 医院、A 医院、D 医院。就综合绩效水平而言四所样本医院的相对接近度则为 0.472~0.581，医院之间差异明显，从相对接近度得出医疗服务绩效水平最高是 C 医院，D 医院则最低。基于 TOPSIS 法的医疗服务绩效评价结果见表 4、表 5。

2.4.4 三种评价结果相关性分析

从三种综合评价方法计算的排序结果来说，方法之间存在不同程度上的差异，需要利用相关分析去研究三种综合评价方法之间的相关关系，一般用 Spearman 等级相关系数表示相关关系的强弱情况^[22]。就综合绩效而言，基于 DRG 的医疗服务绩效评价结果表明三种排序结果具有良好的一致性，有密切相关关系（ $P < 0.05$ ），说明三种综合评价方法能够科学客观地反映四所样本医院的真实绩效水平，为评价院际间的绩效水平提供了数据支持。具体 Spearman 等级相关系数检验结果见表 6。

表 6 医疗服务综合绩效评价结果的 Spearman 等级相关关系检验

评价方法	综合指数法	RSR 法	TOPSIS 法
综合指数法	1.000 ($P < 0.010.000^{***}$)	0.729($P < 0.010.000^{**}$) ($P < 0.010.000^{**}$)	0.669($P < 0.010.001^{*}$) ($P < 0.010.001^{*}$)
RSR 法	0.729($P < 0.010.000^{**}$) ($P < 0.010.000^{**}$)	1.000 ($P < 0.010.000^{***}$)	0.910($P < 0.010.000^{*}$) ($P < 0.010.000^{*}$)
TOPSIS 法	0.669($P < 0.010.001^{*}$) ($P < 0.010.001^{*}$)	0.910($P < 0.010.000^{*}$) ($P < 0.010.000^{*}$)	1.000 ($P < 0.010.000^{***}$)

注：***、**、*分别代表1%、5%、10%的显著性水平

3.讨论

3.1 医疗服务绩效评价**指标体系**指标的构建突出价值医疗导向

当前医院排行榜及学者构建的绩效评价体系多侧重于声誉、规模或单一维度，存在对患者体验、医疗质量安全等关键指标关注不足，本研究围绕 DRG 支付方式改革，深度融合价值医疗理念，构建三级医院医疗服务绩效评价**指标体系**，除医疗服务能力、医疗服务效率等常规一级指标，还增加医保结算清单质量、患者体验、持续发展等一级指标，在指标选取上，坚持宏观与微观、高优与低优、客观与主观相结合，既包含 CMI、DRG 组数等反映整体能力的宏观指标，也纳入再入院率、死亡率等微观质量指标，并首次整合了患者满意度、个人自付比例等体现“以患者为中心”的指标，彰显了价值医疗的理念。

3.2 评价指标**体系优化**权重的**合理化分配**，**平衡主客观因素**

科学合理的指标权重确定是绩效评价体系的核心环节，其准确性直接决定了评价结果的客观性和有效性^[23]。以往研究在确定医疗服务绩效评价**指标权重**时，一般都会选择主观赋权法或客观赋权法^[24]。本研究采用的博弈论法多用于分析多个决策主体行为相互影响时的理性行为及其决策均衡的问题，当博弈的各方在决策过程中为了使共同的利益实现最大化，通常会做出妥协，旨在缩小各决策方权重与组合权重之间的偏差之和^[25]。这一赋权法平衡了专家主观判断和实际客观数据，增强了权重的科学性与引导性，使确定的指标权重更加合理。

3.3 多种综合评价方法的应用**基于多种综合评价方法的绩效评价策略**

目前国内外大部分文献已由单一指标向综合评价方法扩展，但仍采用一种综合评价方法评价医疗服务绩效水平^{[26][27]}，这种比较单一的评价模式下可能会导致评价结果不够全面。本研究在进行绩效评价时选择了三种综合评价方法，并经 Spearman 等级相关系数分析相关关系，弥补了单一评价方法可靠性不强的局限性，一方面使绩效评价结果兼具全面性与客观性，另一方面也形成了一套横向比较的基准，用于分析各医院在不同维度存在的绩效水平差异，通过横向比较促使各医院明确其在发展中的优劣势，并制定改进措施。

3.4 样本医院**基于指标体系**的绩效评价结果及不同综合评价方法的差异分析

分维度进行绩效评价和时，发现三种评价结果的排序存在差异，但大部分评价结果相近，表明三种综合评价方法可靠，结果稳定。进行综合绩效评价时结果显示，综合指数法所得结果中排在首位与末位的分别为 D 医院与 C 医院，RSR 法与 TOPSIS 法恰巧相反，其与组合评价结果一致，首位、末位分别为 C 医院与 D 医院。具体来说 B、C 两所医院作为省内顶尖的三甲综合医院综合实力上各有优势，在医保支付改革背景下的适应能力和管理

水平尚存在一定的差异。C 医院凭借医疗服务能力和效率维度的突出表现，在 RSR 法和 TOPSIS 法中综合排名第一；而 B 医院则在医疗服务能力和患者体验方面表现优异，但受医疗质量安全等维度短板影响，其综合排名略逊于 C 医院；此外，A 医院整体表现相对来说比较稳定，提示其仍有进一步的提升空间。不同评价方法的综合绩效水平差异可能源于评价方法的侧重点差异，如果数据存在偏态分布、异常值或指标间高度相关性，指标值没有上下限的约束，综合指数法可能放大某些指标的作用^[28]，D 医院作为专科医院位居第十一名（全国 84 家相关医院），其虽在相同性质的医疗机构中处于前列，专业优势明显，但是与综合医院相比，其受限于专科医院的性质及服务范围的狭窄性，故在 RSR 法和 TOPSIS 法评价下综合绩效水平较低，有待进一步提升。RSR 法和 TOPSIS 法更加强调多维均衡性，能更全面地反映整体表现，也暴露出 D 医院因专科性质和服务范围局限导致的患者体验、医疗服务能力等综合能力短板，这一结果提示，专科医院需在保持专业优势的同时，警惕评价体系差异对绩效定位的影响，针对性提升综合服务能力以实现全面发展。

4 建议

4.1 医院层面：立足绩效评价结果推进分类改进与内涵建设立足样本医院实际，推进精准化改进提升

结合四所样本医院的绩效评价结果，立足各医院功能定位与发展短板，实施差异化、精准化的靶向改进，依托指标体系的导向作用，驱动医院管理效能与服务质量持续提升，同时进一步验证指标体系的实践指导价值。如四所样本医院要结合自身功能定位，持续优化管理策略，基于各维度评价发现的优势与短板，实施差异化、精准化的靶向改进，驱动医院管理效能与服务质量持续提升。C 医院在医疗服务能力与效率维度表现突出，综合排名靠前，建议其进一步巩固优势，将高效率服务模式标准化，并重点加强医疗质量安全维度的精细化管理，形成全面均衡的发展格局。B 医院在患者体验与医疗服务能力方面表现优异，但受质量安全等维度影响综合排名，需针对性加强临床路径管理与质量控制体系建设，弥补短板。A 医院表现总体稳定但提升空间明显，应借鉴排名靠前医院的优势维度经验，重点优化运营流程与资源分配效率。D 医院作为专科医院，虽在专业领域优势显著，但在患者体验、综合服务能力等方面存在不足，需在保持专科特色的同时，拓展服务维度优化服务流程，增强多学科协作与综合服务能力，避免因评价体系不同导致的绩效定位偏差。

4.2 基于实践反馈优化指标体系，强化验证与适配

四所样本医院的实证结果为指标体系的完善提供了直接依据，结合样本医院实际应用效果，为提升指标体系的科学性、普适性与导向性可以从以下几方面优化：一是要细化指标设置，兼顾共性与个性，保留医疗服务能力等核心共性指标，保障不同类型医院横向对

比的合理性；针对专科医院增设专科特色指标，契合其“专病专治”定位，为综合医院细化 DRG 相关指标，适配其多元诊疗需求。二是建立权重动态调整机制，每 1~2 年结合样本医院实践数据与专家咨询意见修订指标权重，有效减少主观判断与数据敏感性对评价结果的影响，确保权重分配与医疗服务发展实际同频。三是扩大指标体系验证范围，覆盖其他不同地区的三级医院，持续优化指标定义与核算标准；同时加强与“国考”绩效考核指标的衔接，进一步验证并提升指标体系在不同情境下的可靠性与实践价值，充分发挥其绩效导向作用。

4.3 健全多主体协同机制，优化指标体系构建外部支撑

为夯实指标体系应用基础，需健全医保部门、患者、卫生行政部门多主体协同机制。医保支付方面，依托结算数据反向验证、优化指标体系，将服务效率、质量安全等评价结果作为医保支付政策调整的参考依据，检验指标区分度与敏感性；结合医疗技术进步与疾病谱变化动态优化 DRG 分组方案，从实际出发修订、完善 DRG 分组及权重，使其更加契合临床实际情况。患者参与方面，要强化患者参与，引导其查阅电子健康档案、监督诊疗配合随访康复，以患者反馈倒逼服务提升，支撑患者体验维度指标完善。此外，针对学科骨干数据因统计口径不统一无法纳入评价的问题，提示卫生行政部门可以制定统一的统计标准，为各医院收集相关数据提供明确的指导，这能反映出医院在学科建设及人才培养方面的成绩，也有助于完善医疗服务绩效评价指标体系，从而更全面地评价医疗服务绩效水平。

4.2 医保层面：完善 DRG 支付机制与动态调整体系

医保部门应坚持国家政策导向，探索出适合本省医院的 DRG 支付方式改革路径，并在省内符合条件的所有医院中覆盖 DRG 支付方式，形成医保支付新机制。结合各医院绩效评价中反映出的差异，实施分类引导与动态调整，如对 C、B 等综合医院可进一步推动基于效率与质量的激励支付机制；对 D 等专科医院，可探索符合其专科特点的 DRG 分组与权重设置，避免“一刀切”。同时，需建立分组与权重的定期修订机制，结合疾病谱变化与医疗技术进步，增强 DRG 制度的临床适应性与公平性。

4.3 患者层面：强化主动参与意识与医患协同治理

患者要积极行动来消除信息不对称带来的负向影响，作为直接接受医疗服务的主体，其可以主动获取就诊过程中的电子健康档案，关注检查、化验等结果，以动态监督的方式推动诊疗过程透明化，从而能正向反馈医疗服务质量。另一方面，患者还需要增强就医过程中的依从性，配合出院后的随访工作及康复管理工作，在全周期的健康管理中实现与医

院间的责任共担，既能提升其自身的就医体验，还能为持续改进医疗服务贡献力量。

作者贡献：马佳妮、马亦晨负责文章的构思、设计与撰写，资料收集与分析；韩雪梅负责文章思路指导和研究整体设计；张雨佳、万婉参与图表制作和论文后期修改完善；刘艳英负责论文审阅与修订，以及论文整体质量控制。

作者声明本文无实际或潜在的利益冲突

参考文献：

- [1] 谢春艳,王海银,王美凤,等.基于价值的战略购买:健康中国背景下医保治理的现代化转向[J].中国卫生事业管理, 2022, 39(05): 348-352.
- [2] Porter M E, Teisberg E O. Redefining health care: creating value-based competition on results [M]. Harvard: Harvard Business Press, 2006.
- [3] Busse R, Geissler A, Aaviksoo A, et al. Diagnosis related groups in Europe: moving towards transparency, efficiency, and quality in hospitals? [J]. BMJ, 2013, 346: f3197.
- [4] Chen Y J, Zhang X Y, Yan J Q, et al. Impact of Diagnosis-Related Groups on Inpatient Quality of Health Care: A Systematic Review and Meta-Analysis [J]. Inquiry the Journal of Health Care Organization Provision and Financing, 2023, 60-75.
- [5] Roger France F H. Case mix use in 25 countries: a migration success but international comparisons failure [J]. International Journal of Medical Informatics, 2003, 70(2-3): 215-219.
- [6] 刘超.德尔菲法起源及其在生物医学领域应用中的研究[D].太原:山西医科大学, 2021.
- [7] 周硕,芮国忠.新药研发项目风险评估方法的应用[J].中国医药技术经济与管理, 2008, (05): 61-72.
- [8] 胡燕红,和红燕.大学生思政课获得感评价指标体系构建——基于德尔菲法和层次分析法[J].河南科技学院学报, 2024, 44(12): 1-11.
- [9] 刘秋艳,吴新年.多要素评价中指标权重的确定方法评述[J].知识管理论坛, 2017, 2(06): 500-510.
- [10] 杨宇.多指标综合评价中赋权方法评析[J].统计与决策, 2006, (13): 17-19.
- [11] 张立军,张潇.基于改进 CRITIC 法的加权聚类方法[J].统计与决策, 2015, (22): 65-68.
- [12] 陈加良.基于博弈论的组合赋权评价方法研究[J].福建电脑, 2003, (09): 15-16.
- [13] 接纯纯,周典,田帝,等.基于 DRG 和综合指数法的住院患者医疗服务绩效评价[J].中国医院管理, 2022, 42(10): 27-30.
- [14] 魏洁群,王小燕,谢玉婷,等.基于综合指数法的某三甲医院神经外科 DRG 绩效评价研究[J].新疆医学, 2024, 54(08): 1035-1038.
- [15] 郭印.天津市社区卫生服务机构的绩效研究[D].天津:天津大学, 2009.
- [16] 胡晶,张会娜,王宏,等.综合评价法在中医药临床疗效评价中的应用[J].北京中医药, 2023, 42(05): 482-486.
- [17] 牛文奇.基于 DRG 导向的三级综合医院绩效评价指标体系的构建研究[D].青岛:青岛大学, 2022.
- [18] 滕佳利,宋桂杭,李倩,等.基于 DIP 的医保服务质量评价指标体系的构建与应用——基于德尔菲法和层次分析法[J].中国卫生政策研究, 2023, 16(02): 29-35.

- [19] 张振.天津市 DRG 医保支付方式改革实施效果评价研究[D]天津:天津工业大学, 2023.
- [20] 沈绮云,欧阳河,欧阳育良.产教融合目标达成度评价指标体系构建——基于德尔菲法和层次分析法的研究[J].高教探索, 2021, (12): 104-109.
- [21] 吴谦,邱映贵.国内外医院医疗质量评价方法分析研究[J].中国医院, 2019, 23(10): 25-28.
- [22] 王婧.县级公立医院妇科服务能力评价研究[D]武汉:华中科技大学, 2023.
- [23] 范涌峰,宋乃庆.大数据时代的教育测评模型及其范式构建[J].中国社会科学, 2019, 40(12): 139-155.
- [24] 罗荣,金曦,杨琦,等.妇幼保健机构综合绩效评估指标体系权重研究[J].中国卫生统计,2009,26(02):173-175.
- [25] 李洪波,王成文,吴瑞,等.基于博弈论法的银川市地下水资源承载力评价[J].人民黄河,2023,45(S1):42-43.
- [26] 陈世景,谢海梅,林丽丽,等.基于 DRGs 数据的 TOPSIS 方法对住院服务质量评价[J].中国病案, 2017, 18(02): 37-39.
- [27] 陈甜甜,陈霞,阎景红,等.秩和比法结合疾病诊断分组在护理绩效管理中的应用[J].中国病案, 2021, 22(06): 19-22.
- [28] 陈锦华.叠代法、综合指数法、TOPSIS 法在评价某医院综合效益中应用的比较[J].中国医院统计, 2003, (01): 19-22.