

基于未确知测度理论的多科性医院医疗 大数据治理能力成熟度评价

谢刚¹, 孙玉军², 李月云³

(1. 江苏大学管理学院, 镇江 212013; 2. 苏州大学附属第一医院, 江苏 苏州 215006;
3. 江苏大学理学院, 江苏 镇江 212013)

摘要:为实现对多科性医院大数据治理水平的分析和评估,在数据管理成熟度模型和能力成熟度模型基础上,构建多科性医院大数据治理能力成熟度评价模型。将医院大数据治理能力成熟度分为初始级、程序级、规范级、管理级和优化级五个等级,基于医院内部大数据治理特征,建立包括三级指标的成熟度评价指标体系。然后,分别用客观赋权法及组合赋权法确定各指标权重,并结合未确知测度理论和评价方法,量化评估两个三甲多科性医院大数据治理能力的成熟度。结果表明,两种赋权方法下两家医院的大数据治理能力水平均处于规范级,与两家医院大数据治理的实际情况相吻合。未确知测度理论与信息熵权相结合,可以平衡指标测量主观性和评价结果稳定性,建立的模型能为多科性医院大数据治理提供具有参考价值的决策信息。

关键词:医疗大数据;大数据治理;能力成熟度评价

中图分类号:G203 **文献标志码:**A **文章编号:**1002-980X(2019)09-0089-08

1 研究背景

大数据时代,教育、科技、电商领域的数据“潘多拉魔盒”逐渐被打开。医疗领域,大数据也越来越多地应用于精准医疗。利用海量数据进行产品和服务创新,优化流程、产品和服务,离不开大数据治理。目前,多科性医院内部普遍信息系统板块较多,各个信息系统累积下来的医疗数据不仅体量大,还类型繁多、来源多样,既有来自患者的健康数据,又有来自患者与医生、医生与医生之间的交互数据,以及患者支付的交易数据和各种穿戴设备(如手环、起搏器、眼镜等)收集到的观测数据。这些数据以不同的格式、形式孤立地存在于各个系统中。许多多科性医院对医疗大数据的应用能力不够,缺乏有效的数据治理机制,在数据应用过程中面临异构数据储存、处理、融合、共享和安全方面的诸多挑战。这对有效利用医疗大数据优化医疗药品、流程和服务造成了很大的困扰。因此,建立多科性医院大数据治理能

力成熟度评价指标体系,提高医疗大数据治理水平,显得十分迫切。

大数据治理是广义信息治理计划的一部分,即制定与大数据有关的数据优化、隐私保护与数据变现的政策^[1]。大数据治理需将大数据整合到既有的信息治理框架中,并且将大数据与元数据、隐私、数据质量和元数据等信息治理准则结合^[2-3]。医疗领域大数据治理引起了研究者的关注。例如,医疗数据获取的隐私和安全^[4-5],数据的所有权问题^[4],健康大数据应用于精准医疗面临的技术壁垒^[6],数据标准和质量控制^[5-7],数据集建立与管理^[8-10]等方面是比较热点的领域。这些研究对于理解数据治理的问题和特征具有重要价值,但大多数侧重于分析医疗大数据管理活动和知识,鲜有探索与评估多科性医院医疗大数据治理活动的过程和实践的发展水平。

能力成熟度模型最初是一套基于软件开发过程的管理、能力提高的评估模式和方法。近年来,它被

收稿日期:2019-07-24

基金项目:江苏省高校哲学社会科学基金项目“基于医疗大数据的智慧医疗商业模式构建与对策研究”(2016SJB630095);江苏大学高级人才科研基金项目“社会化媒体环境下基于‘转推’模式的营销信息传播机制”(14JDG105)

作者简介:谢刚(1974—),男,四川广安人,江苏大学管理学院,副教授,博士,研究方向:信息管理、技术创新;孙玉军(1973—),男,安徽来安人,苏州大学附属第一医院,高级审计师,硕士,研究方向:医院管理与内部控制;李月云(1998—),女,江苏泰州人,江苏大学理学院,本科生,研究方向:评价理论与方法。

应用于数据管理,例如美国雪城大学的研究数据管理能力成熟度模型^[11-12]。另一方面,关于数据管理、数据治理的成熟度测评模型的研究不断拓展。Ryu 等从数据质量角度考虑了元数据与数据结构,提出四级成熟度模型,用以评估和管理企业数据质量等^[13]。王刚等从数据支撑管理、数据治理惯例、数据服务能力等层面,采用证据分组的方法对银行数据治理进行评价^[14]。张宇杰等从战略规划、组织和制度保障、技术架构、数据管理和治理能力等层面构建政府大数据治理的成熟度测评指标体系^[15]。还有党洪莉等建立的数据管理成熟度模型^[16]。这些数据治理成熟度模型的对象主要为企业数据、高校研究数据和政府数据,针对多科性医院特征的医疗大数据较少。

现有的成熟度评价常采用模糊综合评价法。但这种方法较重视主观因素,忽视客观因素,使得评价结果准确性不高。未确知测度理论是在未确知信息(即决策时必须利用但不能确知的信息)基础上建立起来的解决综合评价问题的理论和方法。未确知测度理论是一种以模糊推理为主的定量与定性相结合的方法,被应用于能力评价^[17-21]。多科性医院医疗大数据多源异构,大数据治理活动环节交叉关联,医院大数据治理能力信息具有随机模糊性,各指标间相互关联制约。采用未确知测度理论评价多科性医院医疗大数据治理能力成熟度,可以建立多个评价对象间的相关关系,并有利于处理未确知信息,提高评价结果的可靠性。本文基于未确知测度理论,提出多科性医院大数据治理能力成熟度,构建能力成熟度评价模型和方法,为多科性医院的大数据治理改进和评估提供一个有效的方法和工具。运用成熟度模型对多科性医院大数据治理水平进行评估,提供改进路径,有助于多科性医院提高大数据治理能力。

2 多科性医院大数据治理能力成熟度等级

大数据治理的活动包括数据治理层面的治理架构、治理政策和数据标准;还包括运作层面的数据管理工具、方法,数据应用方面的流程、人员配置、数据业务管理等^[22]。通过构建多科性医院大数据治理能力成熟度模型,可以让医院明确应该进行哪些数据管理活动、每个数据管理活动会经历哪些过程,目前医院的数据治理处于何种层次,为提高治理层次需做出怎样的改进。基于现有的数据治理成熟度模型,本文从大数据治理活动及多科性医院内部大数据特点出发,将多科性医院大数据治理能力成熟度划分为 5 个等级,如图 1 所示。

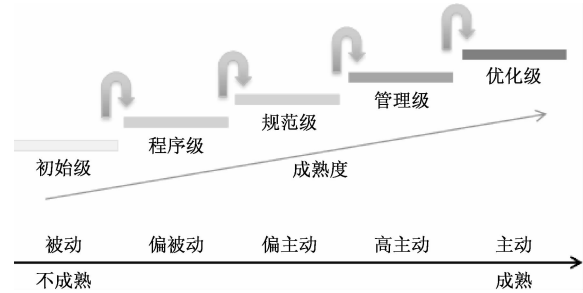


图 1 多科性医院大数据治理能力成熟度等级

(1)初始级:处于这一阶段的医院大数据治理水平的典型特征为:医院没有正确认识到医疗大数据治理的重要性,因而对其关注程度较低。一方面,医院对于大数据治理的责任不明确且各部门没有确切的任务分工。另一方面,医院既没有建立标准的大数据治理体系,也没有健全的大数据治理管理制度。此时,医疗大数据治理水平还处于一种比较粗犷的状态。

(2)程序级:随着长时间的发展与累积,医院逐渐意识到医疗大数据的有效治理对提高医疗服务的重要性。医院开始逐步完善大数据治理的基础设施,并建立大数据治理的基本标准化体系以及管理组织与制度,在规范化各种医疗数据的同时对大数据处理进行实时监控。此阶段医院开始构建数据治理的组织结构,并且注重培养医护人员的专业素养。

(3)规范级:程序级注重从静态的角度制定治理标准、治理组织和管理制度等,但在实际的治理过程中,环境和患者需求的动态变化会对医院大数据治理提出更高的要求,因而,规范级需要从静态设计向动态环境进行适应与改变。医院管理者需对医护人员进行大数据治理的认知教育,使其不断更新自己的知识体系以适应新的环境。此外,医院也在招聘与培养大数据治理专业人才加以辅助。该阶段医院数据治理制度、数据管理流程和数据标准化已达到较为规范化的水平。

(4)管理级:医院对医疗数据的管理与利用水平显著提高,医院大数据治理收到预期效果。医院不仅对医疗大数据进行实时更新,还对医疗健康大数据进行隐私保护,不断提高基于大数据的医疗服务能力,患者治疗成本降低的同时,也提高了医护人员的工作效率。

(5)优化级:第五级是多科性医院大数据治理能力成熟度的最高级别,此阶段医院大数据治理能及时响应医院内部与外部的医疗政策,基于医疗大数据的服务能力已达到较高水平。医院通过不断改进优化自己的大数据治理措施来更好地服务社会,为患者谋福利。

3 医院大数据治理能力成熟度评价指标体系构建

根据构建综合评价指标体系的主要原则^[23],结合专家意见,并考虑多科性医院大数据治理的工作

环境,将医院大数据治理能力成熟度评价指标体系分为三个层次,其中一级指标包括支撑条件、管理基础和动态适应三项;二级指标包括组织、专业人才和基础措施等 7 项;三级指标包括数据治理组织结构与权限等 22 项。指标体系详如表 1 所示。

表 1 多科性医院大数据治理能力成熟度评价指标体系

目标	一级指标	二级指标	三级指标	
多科性医院 大数据治理 能力成熟度	支撑条件 S_1	组织 X_1	数据治理组织结构与权限 X_{11}	
			数据治理组织流程 X_{12}	
		专业人才 X_2	大数据架构师 X_{21}	
			大数据资产管理专业人才 X_{22}	
	管理基础 S_2	基础设施 X_3	医护人员的数据素养 X_{23}	
			数据治理的基础设施 X_{31}	
		治理规划与制度 X_4	大数据云计算环境下安全检测与防护技术 X_{32}	
			元数据及数据标准 X_{41}	
		数据质量 X_5	大数据架构规划 X_{42}	
			大数据治理的管理制度 X_{43}	
			数据来源维度多样性 X_{51}	
			元数据一致性 X_{52}	
			多源异构数据归类与融合 X_{53}	
			诊疗和健康数据采集完整性 X_{54}	
		动态适应 S_3	基于大数据的医疗服务能力 X_6	诊疗和健康数据更新时效性 X_{55}
				医疗健康数据隐私保护 X_{56}
				基于大数据的医疗服务响应 X_{61}
			外部响应能力 X_7	大数据于医疗业务覆盖程度 X_{62}
	基于大数据的医疗服务需求匹配 X_{63}			
	响应政府大数据政策的能力 X_{71}			
			对社会舆论的应对能力 X_{72}	
			响应国家医疗改革的能力 X_{73}	

4 医院大数据治理能力成熟度评价方法

本文采用未确知测度模型来判断三甲多科性医院的大数据治理成熟度等级。首先,通过未确知测度概念构造单指标测度函数与测度评价矩阵。其次,通过客观赋权法(信息熵)、主客观组合赋权法等方法分别确定各项指标的权重后,再通过测度矩阵与权重向量计算多指标综合测度。最后,利用置信度识别判断医院大数据治理能力处于哪类成熟度等级。

设现有评价对象 n 个,每个评价对象有 m 个评价指标,将成熟度等级分为 p 类,其相应的数学符号的解释如表 2 所示。

若多科性医院大数据治理成熟度等级逐步增高,则有 $C_1 < C_2 < \dots < C_p$;若成熟度等级逐步降低,则有 $C_1 > C_2 > \dots > C_p$ 。称 $\{C_1, C_2, \dots, C_p\}$ 是属于评价等级空间的有序分割类。

4.1 单指标测度函数构造

现将 x_{ij} 记作评价对象 R_i 关于评价指标 I_j 的测量值,并将 x_{ij} 划分为 p 个等级,得到评价等级空间

表 2 数学符号及其解释

模型符号	解释
n	评价对象个数
m	评价指标个数
p	评价等级个数
R_i	第 i 个评价对象
I_j	第 j 个评价指标
C_k	第 k 个评价等级
ω_j	信息熵法第 j 个评价指标的权重
ω_j^*	组合赋权得到的第 j 个评价指标的权重
u_j	G1 法第 j 个评价指标的权重
u_j^*	G1 法排序后第 j 个评价指标的权重
x_{ij}	评价对象 R_i 关于第 i 个评价指标的测量值
μ_{ijk}	x_{ij} 属于第 k 个评价等级的程度
$R = \{R_1, R_2, \dots, R_n\}$	评价对象空间
$X = \{X_1, X_2, \dots, X_m\}$	评价指标空间
$C = \{C_1, C_2, \dots, C_p\}$	评价等级空间
$W = \{\omega_1, \omega_2, \dots, \omega_m\}$	指标权重空间

$C = \{C_1, C_2, \dots, C_p\}$ 。根据划分等级构造单指标测度函数,通过测度函数求出评价指标测量值 x_{ij} 关于各评价等级的隶属度 μ_{ijk} ,其中:

$$0 \leq \mu_{ijk} \leq 1; \quad (1)$$

$$\sum_{k=1}^p \mu_{ijk} = 1; \quad (2)$$

$$\mu[x_{ij} \in \cup_{k=1}^r C_k] = \sum_{k=1}^r \mu_{ijk}, r = 1, 2, \dots, p. \quad (3)$$

将 μ 称为未确知测度, 简称测度。式(1)、式(2)、式(3)分别称为非负有界性、归一性和可加性。隶属度 μ_{ijk} 构成的矩阵称为单指标测度评价矩阵, 即:

$$B = \begin{bmatrix} \mu_{i11} & \mu_{i12} & \dots & \mu_{i1p} \\ \mu_{i21} & \mu_{i22} & \dots & \mu_{i2p} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{im1} & \mu_{im2} & \dots & \mu_{imp} \end{bmatrix}, i = 1, 2, \dots, n. \quad (4)$$

评价对象 R_i 有 m 个评价指标, 评价等级有 p 个, 因而 B 为 m 行 p 列的矩阵, 且每行数值相加等于 1。

4.2 指标权重确定

要对多科性医院大数据治理能力成熟度进行综合评价, 需要求出各项指标的权重。常用的权重确定方法为主观赋权法、客观赋权法以及组合赋权法。主观赋权法主要有层次分析法、网络层次分析法以及 G1 法等, 它反映决策者的偏好, 主观性偏强; 客观赋权法则包括信息熵法、主成分分析法等, 它过于依赖客观数据, 但忽略了专家的知识 and 经验。由于本文评价方法是基于模糊推理的未确知测度理论, 权重确定若采取主观赋值法, 会加重主观因素的影响, 因此只选用客观赋权法及组合赋权法进行计算。

4.2.1 客观赋权法(信息熵)

权重表示各评价指标影响评价结果的程度, 这里运用信息熵理论, 计算多科性医院大数据治理能力成熟度评价指标体系中各项指标所占的权重^[24]:

$$\nu_j = 1 + \frac{1}{\ln p} \sum_{k=1}^p \mu_{jk} \ln \mu_{jk}; \quad (5)$$

$$\omega_j = \frac{\nu_j}{\sum_{g=1}^m \nu_g}. \quad (6)$$

将 ν_j 称作评价指标 I_j 的熵值, ω_j 称作 I_j 对于评价结果的重要程度, 即为权重。其中, $0 \leq \omega_j \leq 1, \sum_{j=1}^m \omega_j = 1$ 。

4.2.2 组合赋权法

组合赋权法是先利用 G1 法对各评价指标进行主观赋权, 再利用信息熵法进行客观赋权, 然后通过主客观权重值的线性组合得到各项指标的最终权重。

1) 基于 G1 法的主观赋值。

基于 G1 法的主观赋权具体思路为: ①专家按照指标的重要程度进行排序; ②专家对相邻两项指标的相对重要程度进行比较判断, 并赋值; ③根据所赋的值按照公式计算得出每项指标的权重^[24]。详细步骤如下:

第一步, 按照指标的重要程度确定序关系。令 I_1, I_2, \dots, I_m 为 m 项评价指标, 若第 i 项评价指标的重要程度大于第 j 项, 则表示为 $I_i > I_j$ 。若对于指标层 I_1, I_2, \dots, I_m 确定了关系式 $I_1^* > I_2^* > \dots > I_m^*$, 则称 I_1, I_2, \dots, I_m 按“>”确定了序关系。

第二步, 给出相邻两项指标 I_{j-1}^* 与 I_j^* 间的相对重要程度。有些问题仅仅给出序关系不能完全解决问题, 因而需要确定它们之间的相对重要程度。令 u_{j-1}^* 与 u_j^* 表示第 $j-1$ 项指标 I_{j-1}^* 与 j 项指标 I_j^* 的权重, 则 $r_j = \frac{u_{j-1}^*}{u_j^*}$ 表示这两项之间的重要程度之比。

第三步, 计算权重 u_j^* 。专家打分 r_j^* 按照表 3 要求, 则

$$u_m^* = \left[1 + \sum_{k=2}^m \prod_{j=k}^m r_j \right]^{-1}; \quad (7)$$

$$u_{j-1}^* = r_j u_j^* (j = 2, \dots, m-1, m). \quad (8)$$

在得到所有 u_j^* 之后便可以得出原各项指标的权重 u_j 。

表 3 相邻指标间的相对重要程度

r_j	说明(X_{j-1} 与 X_j 相对重要程度)
1.0	同等重要
1.1	重要程度介于同等重要与稍微重要之间
1.2	稍微重要
1.3	重要程度介于稍微重要与明显重要之间
1.4	明显重要
1.5	重要程度介于明显重要与强烈重要之间
1.6	强烈重要
1.7	重要程度介于强烈重要与极端重要之间
1.8	极端重要

2) 组合权重的确定。

设:

$$W_j^* = \beta u_j + (1 - \beta) \omega_j. \quad (9)$$

其中: β 为主观权重在组合权重中所占的比例, $(1 - \beta)$ 为客观权重在组合权重中所占的比例。 ω_j 为信息熵法所得到的第 j 项指标的客观权重值, u_j 为 G1 法得到的第 j 项指标的主观权重值。取 $\beta = 1 - \beta$, 即 $\beta = 0.5$ ^[24]。

4.3 多指标综合测度计算

对于评价对象 R_i , 令 μ_{ik} 表示为 R_i 归属于第 k 级成熟度等级 C_k 的程度, 有:

客观赋权时采用:

$$\mu_{ik} = \sum_{j=1}^m \omega_j \mu_{ijk}, i = 1, 2, \dots, n; k = 1, 2, \dots, p. \quad (10)$$

组合赋权时采用:

$$\mu_{ik} = \sum_{j=1}^m \omega_j^* \mu_{ijk}, i = 1, 2, \dots, n; k = 1, 2, \dots, p; \quad (11)$$

$$\mu_i = \{\mu_{i1}, \mu_{i2}, \dots, \mu_{ip}\}. \quad (12)$$

其中: $0 \leq \mu_{ik} \leq 1$, 且 $\sum_{k=1}^p \mu_{ik} = 1$, μ_i 称为多指标综合未确知测度向量。

4.4 置信度识别

为了评估三甲多科性医院大数据治理能力的成熟度等级, 需要引入置信度识别准则, $\{C_1, C_2, \dots, C_p\}$ 是评价等级空间的一个有序分割类。

设 λ 为置信度(一般 $\lambda \geq 0.5$, 本文取 $\lambda = 0.7$)。

当 $C_1 > C_2 > \dots > C_p$ 时,

$$k_0 = \min\{k: \sum_{l=1}^k \mu_{il} \geq \lambda, k = 1, 2, \dots, p\}; \quad (13)$$

当 $C_1 < C_2 < \dots < C_p$ 时,

$$k_0 = \max\{k: \sum_{l=k}^p \mu_{il} \geq \lambda, k = 1, 2, \dots, p\}. \quad (14)$$

则认为评价对象 R_i 属于第 k_0 级成熟度等级 C_{k_0} 。

5 实证分析

根据 22 项评价指标设计医院大数据治理调查问卷, 采用五分制评分, 成熟度评价等级标准如表 4。选取苏州市和镇江市两家多科性医院发放问卷。苏州某医院问卷调查对象有医护专业人员 60 名、行政管理人员 35 名、信息及数据管理人员 2 名。共发放问卷 97 份, 将选项一致的无效问卷去除, 实际回收有效问卷为 72 份, 对有效数据进行处理, 分析结果如表 5 所示。

表 4 成熟度评价指标划分标准

成熟度等级	初始级	程序级	规范级	管理级	优化级
评价等级	C_1	C_2	C_3	C_4	C_5
分值	0~2	2~3	3~3.75	3.75~4.5	4.5~5

表 5 苏州某医院调查结果

指标	分值	指标	分值	指标	分值	指标	分值
X_{11}	3.389	X_{32}	3.389	X_{53}	3.264	X_{63}	3.431
X_{12}	3.264	X_{41}	3.375	X_{54}	3.500	X_{71}	3.528
X_{21}	3.306	X_{42}	3.278	X_{55}	3.486	X_{72}	3.472
X_{22}	3.250	X_{43}	3.375	X_{56}	3.500	X_{73}	3.722
X_{23}	3.417	X_{51}	3.431	X_{61}	3.375		
X_{31}	3.306	X_{52}	3.375	X_{62}	3.542		

在镇江某医院调查了医护专业人员 90 名、行政管理人员 12 名、信息及数据管理人员 3 名。共发放

问卷 105 份, 实际收回有效问卷 70 份。数据分析结果如表 6 所示。

表 6 镇江某医院问卷调查结果

指标	分值	指标	分值	指标	分值	指标	分值
X_{11}	3.580	X_{32}	3.507	X_{53}	3.507	X_{63}	3.710
X_{12}	3.449	X_{41}	3.696	X_{54}	3.739	X_{71}	3.754
X_{21}	3.551	X_{42}	3.435	X_{55}	3.797	X_{72}	3.754
X_{22}	3.493	X_{43}	3.652	X_{56}	3.696	X_{73}	3.870
X_{23}	3.464	X_{51}	3.667	X_{61}	3.739		
X_{31}	3.348	X_{52}	3.681	X_{62}	3.754		

5.1 单指标测度函数构造

根据成熟度评价指标划分标准(如表 4)构造单指标测度函数, 如图 2 所示。

分别将两家医院的调查结果代入单指标测度函数, 计算得出苏州某医院与镇江某医院的单指标测度评价矩阵。

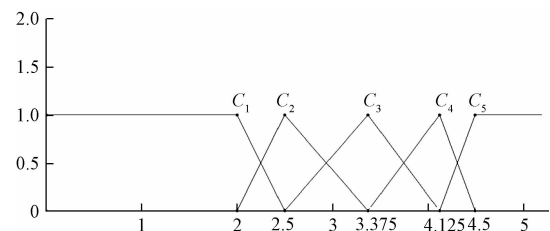


图 2 单指标测度函数

$$B_1 = \begin{pmatrix} 0 & 0 & 0.981 & 0.019 & 0 \\ 0 & 0.127 & 0.873 & 0 & 0 \\ 0 & 0.079 & 0.921 & 0 & 0 \\ 0 & 0.143 & 0.857 & 0 & 0 \\ 0 & 0 & 0.944 & 0.056 & 0 \\ 0 & 0.079 & 0.921 & 0 & 0 \\ 0 & 0 & 0.981 & 0.019 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0.111 & 0.889 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0.925 & 0.075 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0.127 & 0.873 & 0 & 0 \\ 0 & 0 & 0.833 & 0.167 & 0 \\ 0 & 0 & 0.852 & 0.148 & 0 \\ 0 & 0 & 0.833 & 0.167 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0.777 & 0.223 & 0 \\ 0 & 0 & 0.925 & 0.075 & 0 \\ 0 & 0 & 0.796 & 0.204 & 0 \\ 0 & 0 & 0.871 & 0.129 & 0 \\ 0 & 0 & 0.537 & 0.463 & 0 \end{pmatrix},$$

$$B_2 = \begin{pmatrix} 0 & 0 & 0.727 & 0.273 & 0 \\ 0 & 0 & 0.901 & 0.099 & 0 \\ 0 & 0 & 0.765 & 0.235 & 0 \\ 0 & 0 & 0.843 & 0.157 & 0 \\ 0 & 0 & 0.881 & 0.119 & 0 \\ 0 & 0.031 & 0.969 & 0 & 0 \\ 0 & 0 & 0.824 & 0.176 & 0 \\ 0 & 0 & 0.572 & 0.428 & 0 \\ 0 & 0 & 0.920 & 0.080 & 0 \\ 0 & 0 & 0.631 & 0.369 & 0 \\ 0 & 0 & 0.611 & 0.389 & 0 \\ 0 & 0 & 0.592 & 0.408 & 0 \\ 0 & 0 & 0.824 & 0.176 & 0 \\ 0 & 0 & 0.515 & 0.485 & 0 \\ 0 & 0 & 0.437 & 0.563 & 0 \\ 0 & 0 & 0.572 & 0.428 & 0 \\ 0 & 0 & 0.515 & 0.485 & 0 \\ 0 & 0 & 0.495 & 0.505 & 0 \\ 0 & 0 & 0.553 & 0.447 & 0 \\ 0 & 0 & 0.495 & 0.505 & 0 \\ 0 & 0 & 0.495 & 0.505 & 0 \\ 0 & 0 & 0.340 & 0.660 & 0 \end{pmatrix}.$$

5.2 权重确定

5.2.1 客观赋权法

根据式(5)、式(6)以及信息熵理论,利用 matlab 算得苏州某医院和镇江某医院大数据治理能力成熟度评价指标体系中各指标的权重向量:

$$W_1 = (0.0523, 0.0424, 0.0460, 0.0414, 0.0481, 0.0460, 0.0523, 0.0556, 0.0435, 0.0556, 0.0464, 0.0556, 0.0424, 0.0400, 0.0411, 0.0400, 0.0556, 0.0372, 0.0464, 0.0381, 0.0423, 0.0317);$$

$$W_2 = (0.0446, 0.0560, 0.0464, 0.0512, 0.0542, 0.0641, 0.0498, 0.0404, 0.0580, 0.0414, 0.0410, 0.0407, 0.0498, 0.0399, 0.0403, 0.0404, 0.0399, 0.0399, 0.0402, 0.0399, 0.0399, 0.0422).$$

5.2.2 组合赋权法

一级指标对于准则层的计算:根据专家意见,指标重要程度为 $X_1 > X_2 > X_3 > X_4 > X_5 > X_6 > X_7$,且相对重要程度为 $X_1/X_2 = 1.3, X_2/X_3 = 1.1, X_3/X_4 = 1.0, X_4/X_5 = 1.0, X_5/X_6 = 1.1, X_6/X_7 = 1.1$,则 $r_2 = 1.3, r_3 = 1.1, r_4 = 1.0, r_5 = 1.0, r_6 = 1.1, r_7 = 1.1$,根据公式(7)、式(8), $X_1, X_2, X_3, X_4, X_5, X_6, X_7$ 的权重分别为 0.2、0.15、0.14、0.14、0.14、0.12、0.11。

同理可求得二级指标对于所属一级指标的权重,从而求出基于 G1 法的主观权重,将主观权重与客观权重组合得到组合权重,具体数值如表 7 所示。

5.3 多指标综合测度计算

5.3.1 客观赋权下的多指标测度函数

根据式(10),将客观权重与单指标测度评价矩阵相乘,就可以得到多指标综合未确知测度向量,即

$$\mu_1 = \{0, 0.0288, 0.9039, 0.0673, 0\};$$

$$\mu_2 = \{0, 0.0020, 0.6827, 0.3153, 0\}.$$

5.3.2 组合赋权下的多指标测度函数

根据式(11),将组合权重与单指标测度评价矩阵相乘,就可以得到多指标综合未确知测度向量,即

$$\mu_1 = \{0, 0.0341, 0.9042, 0.0617, 0\};$$

$$\mu_2 = \{0, 0.002, 0.696, 0.302, 0\}.$$

5.4 置信度识别

设医院大数据治理成熟度等级逐步增高,所以 $C_1 < C_2 < \dots < C_p$,选择式(14)与多指标综合未确知测度向量 μ 结合,进行置信度识别。

5.4.1 客观赋权下的置信度识别

$$k_{01} = \max\{k: 0.9039 + 0.0673 = 0.9712 \geq 0.7\} = 3;$$

$$k_{02} = \max\{k: 0.6827 + 0.3153 = 0.9980 \geq 0.7\} = 3.$$

表 7 各项指标的主客观权重及组合权重

一级指标	一级指标权重	二级指标	二级指标权重	主观权重	客观权重 (苏州某医院)	客观权重 (镇江某医院)	组合权重 (苏州某医院)	组合权重 (镇江某医院)
X ₁	0.2	X ₁₁	0.52	0.104	0.0523	0.0446	0.0782	0.0743
		X ₁₂	0.48	0.096	0.0424	0.0560	0.0692	0.0760
X ₂	0.15	X ₂₁	0.42	0.063	0.0460	0.0464	0.0545	0.0547
		X ₂₂	0.33	0.0495	0.0414	0.0512	0.0455	0.0504
		X ₂₃	0.25	0.0375	0.0481	0.0542	0.0428	0.0459
X ₃	0.14	X ₃₁	0.52	0.0728	0.0460	0.0641	0.0594	0.0685
		X ₃₂	0.48	0.0672	0.0523	0.0498	0.0598	0.0585
X ₄	0.14	X ₄₁	0.38	0.0532	0.0556	0.0404	0.0544	0.0468
		X ₄₂	0.35	0.049	0.0435	0.0580	0.0463	0.0535
		X ₄₃	0.27	0.0378	0.0556	0.0414	0.0467	0.0396
X ₅	0.14	X ₅₁	0.18	0.0252	0.0464	0.0410	0.0358	0.0331
		X ₅₂	0.28	0.0392	0.0556	0.0407	0.0474	0.0400
		X ₅₃	0.22	0.0308	0.0424	0.0498	0.0366	0.0403
		X ₅₄	0.08	0.0112	0.0400	0.0399	0.0256	0.0256
		X ₅₅	0.08	0.0112	0.0411	0.0403	0.0262	0.0258
		X ₅₆	0.16	0.0224	0.0400	0.0404	0.0312	0.0314
X ₆	0.12	X ₆₁	0.31	0.0372	0.0556	0.0399	0.0464	0.0386
		X ₆₂	0.31	0.0372	0.0372	0.0399	0.0372	0.0386
		X ₆₃	0.38	0.0456	0.0464	0.0402	0.0460	0.0429
X ₇	0.11	X ₇₁	0.35	0.0385	0.0381	0.0399	0.0383	0.0392
		X ₇₂	0.3	0.033	0.0423	0.0399	0.0377	0.0365
		X ₇₃	0.35	0.0385	0.0317	0.0422	0.0351	0.0404

5.4.2 组合赋权下的置信度识别

$$k'_{01} = \max\{k; 0.9042 + 0.0617 = 0.9659 \geq 0.7\} = 3;$$

$$k'_{02} = \max\{k; 0.696 + 0.302 = 0.998 \geq 0.7\} = 3。$$

由分析可知,两种赋权方法下苏州某医院与镇江某医院的大数据治理能力均处于C₃级,即规范级。这与两家医院实际治理水平相吻合。说明两家医院的大数据治理意识较强,采取了较为有效的措施来提升医护人员对医疗大数据治理的认知;数据治理制度、数据管理流程和数据标准化形成了较为规范化的体系;建立了较为完善的数据治理支撑体系,不断招聘和培养大数据管理方面专业人员;能够较好地根据政府、社会的要求适时地调整医院大数据治理架构和管理过程。这些方面均是多科性医院提升大数据治理能力的有效途径。

6 结论

本文根据多科性医院医疗大数据的特征和治理现状,提出5个阶梯状多科性医院大数据治理能力成熟度进化等级及其成熟度评价指标体系,结合实际案例,基于未确知测度理论评价成熟度水平。

(1)构建的能力成熟度模型可以帮助多科性医院理解大数据管理过程、实践活动应该达到一个什么样的发展程度和水平。医院可以准确判断医疗大数据治理水平处于何种成熟度等级,有利于把控大数据治理体系的运行状态,提高治理的主动性。

(2)未确知测度理论与信息熵权相结合可以平衡指标测量主观性和评价结果稳定性之间的矛盾,为医院数据治理管理层提供具有参考价值的决策信息。通过未确知测度理论与多种赋权方法相结合对两家多科性医院大数据治理能力成熟度进行实例评价,结果表明两家医院的能力成熟度均属于规范级,这与实际运行情况相吻合,验证了未确知测度理论与信息熵权相结合的成熟度评价模型及方法的有效性。可以反映两家医院大数据治理水平的总体状况,为多科性医院实施医疗大数据治理的改进方向和措施调整提供指导。

(3)深入研究影响多科性医院大数据治理能力成熟度的因素,探讨提高多科性医院大数据治理能力的途径,将是下一步的研究方向。

参考文献

- [1] 桑尼尔·索雷斯. 大数据治理[M]. 匡斌,译. 北京:清华大学出版社, 2014: 4.
- [2] 童楠楠, 朝乐门. 大数据时代下数据管理理念的变革: 从结果派到过程派[J]. 情报理论与实践, 2017, 40(2): 60-65.
- [3] ABRAHAM R, SCHNEIDER J, BROCKE J. Data governance: a conceptual framework, structured review, and research agenda[J]. International Journal of Information Management, 2019(49): 424-438.
- [4] ROSENBAUM S. Data governance and stewardship: designing data stewardship entities and advancing data ac-

- cess[J]. Health Services Research, 2010, 45(5): 1442-1455.
- [5] 张国庆, 李亦学, 王泽峰, 等. 生物医学大数据发展的新挑战与趋势[J]. 中国科学院院刊, 2018, 33(8): 853-860.
- [6] 武志慧, 王飞, 姜召芸, 等. 健康医疗大数据与罕见病的精准用药[J]. 科技导报, 2017, 35(16): 20-25.
- [7] 吴泰相, 卞兆祥, 李幼平, 等. 促进我国临床试验数据管理规范化[J]. 中国循证医学杂志, 2018, 18(6): 532-537.
- [8] 赵烁, 付强, 刘运喜, 等. 统一数据采集规范的区域性医院感染管理信息体系建设[J]. 中华医院感染学杂志, 2016, 26(6): 1404-1406.
- [9] 赵瑜, 万长秀, 周琼, 等. 中医医院护理管理信息基本数据集构建研究[J]. 护理学杂志, 2017, 32(19): 1-5.
- [10] AISLING M, ALENA B, BARBARA P. Big data governance needs more collective responsibility: the role of harm mitigation in the governance of data use in medicine and beyond[J]. Medical Law Review, 2019, fwz016.
- [11] CROWSTON K, QIN J. A capability maturity model for scientific data management: evidence from the literature[J]. Proceedings of the American Society for Information Science and Technology, 2011, 48(1): 1-9.
- [12] Qin J, Crowston K, Flynn C, et al. A capability maturity model for research data management [EB/OL]. [2014-06-11]. Http: //rdm.ischool.syr.edu/xwiki/bin/view/CMM+FOR+RDM/Web-Home.
- [13] Ryu K S, Park J S, Park J H. A Data quality management maturity model[M]. Daejeon, Korea: IT Services Research Division, ETRI, 2006: 191-204.
- [14] 王刚, 汪杨, 王珏, 等. 基于证据分组合成的企业数据治理评价研究[J]. 系统工程理论与实践, 2016, 36(6): 1505-1516.
- [15] 张宇杰, 安小米, 张国庆. 政府大数据治理的成熟度评测指标体系构建[J]. 情报资料工作, 2018(1): 28-32.
- [16] 党洪莉, 谭海兵. 基于 DMM 的数据管理成熟度模型及在服务评估中的应用[J]. 现代情报, 2017, 37(9): 118-121.
- [17] 曹庆奎, 任向阳, 刘琛. 基于粗集-未确知测度模型的企业技术创新能力评价研究[J]. 系统工程理论与实践, 2006(4): 67-86.
- [18] 宋伟, 郑淑芬, 康正. 基于未确知测度的政府信息能力评价研究[J]. 情报科学, 2010(8): 1238-1241.
- [19] 宋伟, 陈伟. 风险社会下企业信息能力的未确知测度评价研究[J]. 情报杂志, 2010, 29(6): 132-135.
- [20] 王彦泽, 王晓峰. 基于未确知测度模型的秦岭华阳古镇景区暴雨防灾减灾能力评价[J]. 自然灾害学报, 2017, 26(2): 67-80.
- [21] 詹丽, 张小月. 基于未确知测度的旅游应急管理能力评价研究——以湖北恩施自治州为例[J]. 资源开发与市场, 2018(1): 113-117.
- [22] SMITH A M. Seven best practices to boost big data governance efforts[J]. Computer Weekly, 2016: 3-5.
- [23] 顾元勋. 产品架构评估: 分析框架与准则[J]. 技术经济, 2019, 38(3): 31-41.
- [24] 戴文磊, 仲雁秋, 曲毅. 基于组合赋权的区域产业安全评价模型及其实证研究[J]. 技术经济, 2017, 36(11): 72-78.

Capability Maturity Evaluation of Medical Big Data Governance in Multidisciplinary Hospitals Based on Unascertained Measure Theory

Xie Gang¹, Sun Yujun², Li Yueyun¹

(1. Jiangsu University, Zhenjiang Jiangsu 212013, China; 2. The First Affiliated Hospital of Soochow University, Suzhou Jiangsu 215006, China; 3. School of Science, Jiangsu University, Zhenjiang Jiangsu 212013, China)

Abstract: In order to realize the analysis and evaluation of big data governance level in multidisciplinary hospitals, an evaluation model of big data governance capability maturity was built on the basis of data management maturity model and capability maturity model. Hospital capability maturity of big data governance is divided into five levels: initial level, program level, standard level, management level and optimization level. Based on the characteristics of hospital internal big data governance, we established a maturity evaluation index system including three indicators. Then, the objective weighting method and the combined weighting method were used to determine the weight of each index, and the maturity of the big data governance ability of two triple-multidisciplinary hospitals was quantitatively evaluated by combining the unascertained measurement theory and evaluation method. The results showed that the big data governance of the two hospitals both achieve the normative level in the different methods, which is consistent with the actual situation of big data governance of the two hospitals. Because the subjectivity of index measurement and the stability of evaluation results are balanced, the model based the unascertained measurement theory together with information entropy can provide many valued information to multidisciplinary hospitals.

Keywords: medical big data; big data governance; capability maturity assessment