

健康医疗数据共享意愿影响因素研究

郭恺琳 樊敏 刘叶 皮洁 党佳荷 马钰潇
(山西医科大学汾阳学院 汾阳 032200)

[摘要]目的 了解居民对健康医疗数据共享意愿及其影响因素，为激发居民数据共享积极性提供依据。方法 采用随机整群抽样和便利抽样方法，通过在线填写问卷调查不同地区的居民，运用R语言进行卡方检验、Logistics回归分析影响居民共享意愿的因素。结果 共回收829份有效问卷，签署同意书(OR=0.393, CI[0.16,0.859])、数据脱敏(OR=0.681, CI[0.471,1.006])、认为有利于科学研究(OR=0.554, CI[0.402,0.768])、认为有利于健康管理(OR=0.564, CI[0.433,0.742])是共享意愿的影响因素。建议 国家应出台健康医疗大数据法细化数据保护制度；发展隐私安全计算技术，保障数据安全；做好宣传普及工作，建立数据共享信心。

[关键词] 健康医疗大数据；共享意愿；影响因素

Survey on influencing factors of willingness to share health care data

GUO Kailin, FAN Min, LIU Ye, PI Jie, DANG Jiahe, MA Yuxiao, Fenyang College of Shanxi Medical University, Fenyang 032200, China

[Abstract] Objective To understand residents' willingness to share health care data and its influencing factors, so as to provide evidence for stimulating residents' enthusiasm for data sharing. Methods Random cluster sampling and convenience sampling were adopted to investigate the residents in different areas by filling out questionnaires online. Chi-square test and Logistics regression were used to analyze the factors affecting residents' willingness to share. Results A total of 829 valid questionnaires were collected. Signing consent (OR = 0.393, CI [0.16,0.859]), data desensitization (OR = 0.681, CI [0.471,1.006]), "considering that it is beneficial to scientific research" (OR=0.554, CI [0.402,0.768]), "considering that it is beneficial to health management" (OR=0.564, CI [0.433,0.742]) are the influencing factors of sharing intention. Suggestion The state should promulgate the laws on big data in health care to refine the data protection; develop privacy security computing technology to ensure data security; do a good job in publicity and popularization work, and build confidence in data sharing.

[Keywords] Big data in health care; sharing willingness; influencing factors

1 引言

健康医疗数据是指在疾病防治、健康管理等过程中产生的与健康医疗相关的数据^[1]，涵盖了贯穿一个人全生命周期的，与个人健康、医疗服务、疾病防控和养生保健有关的数据^[2]，包括身高、体重、血型、基因、电子病历、电子健康档案、药物服用以及生活中的饮食、运动、睡眠等数据^[3]。

健康医疗大数据是国家重要的基础性战略资源，其共享和开放已成为世界各国的普遍共识。数据共享将健康医疗信息系统和公众健康医疗数据互联融合，推进跨区域、跨机构、跨部门的健康医疗，有利于提高医疗服务效率，有助于增加研究价值、验证研究结果、促进取得新发现，实现个性化健康管理及监测等^[4]。健康医疗数据涉及个人基因、家族史、疾病史、药物史等敏感信息，随着人们隐私安全意识的提高，其对健康医疗数据的共享意愿需要给予关注，以保障健康医疗数据共享进程顺利进行。本文通过对居民进行健康医疗大数据共享意愿的调查，分析影响数据共享的因素，了解居民对数据共享的认识及态度，从而为推动数据共享提供理论支持。

†[收稿日期][作者简介]郭恺琳，本科生。通讯作者：樊敏，副教授。

[基金项目]2020年山西省大学生创新创业训练计划项目“居民健康医疗数据共享意愿调查及影响因素分析”(项目编号：2020787)

2 资料与方法

2.1 研究对象

本研究从2020年12月至2021年2月采用随机整群抽样和便利抽样不同地区的居民作为研究对象，运用线上问卷调查法对居民健康医疗数据的共享意愿进行调查。共回收829份合格问卷。

2.2 调查内容及方法

通过查阅文献总结当前各国健康医疗数据共享现状以及共享策略，归纳影响居民健康医疗数据共享意愿的因素，在问卷星中编制问卷。调查内容包括居民基本情况（包括性别、年龄、学历、共享顾虑），对健康医疗数据的共享要求（包括是否需要数据脱敏、是否需要签署知情同意书、是否需要法律保护），对共享健康医疗数据的价值认知(包括是否有利于科学研究、是否能实现健康管理、是否应当给予物质回报)，是否愿意共享健康医疗数据。将是否愿意共享健康医疗数据作为因变量，进行单因素分析和多因素回归分析。见表1。

表 1 变量及赋值表

变量	变量描述	变量赋值
X1	性别	男=1; 女=2
X2	年龄	18岁以下=1; 18~40岁=2; 41~60岁=3; 60岁以上=4
X3	学历	本科以下=1; 本科及以上=2
X4	顾虑	有=1; 无=2
X5	数据脱敏需求	是=1; 否=2; 不清楚=3
X6	签署同意书需求	是=1; 否=2; 不清楚=3
X7	法律保护需求	是=1; 否=2; 不清楚=3
X8	回报需求	是=1; 否=2; 不清楚=3
X9	利于科学研究	同意=1; 不同意=2; 不清楚=3
X10	利于健康管理	同意=1; 不同意=2; 不清楚=3
Y	居民健康医疗数据共享意愿	愿意=1; 不愿意=0

2.3 统计学处理与分析

将收集的数据录入Excel建立数据库，运用R语言进行统计分析，采用卡方检验进行单因素分析，以健康医疗数据共享意愿为因变量，将单因素分析中有统计学意义的变量纳入方程进行多元线性Logistic回归分析，P<0.05为差异有统计学意义。

3 结果

3.1 基本情况

在829份有效问卷中，女性较多，年龄集中在18-40岁之间，学历较高，见表2。其中84%的被调查者表示愿意共享健康医疗数据。91.07%的被调查者有共享顾虑，其中隐私泄露顾虑的约占63.44%，数据去向和用途顾虑约占30.60%，能否得到回报顾虑仅占4.37%，见图1。在数据敏感程度调查中，84.2%的被调查者认为个人信息较为敏感，其次是病史记录、基因信息、体检记录，见图2。

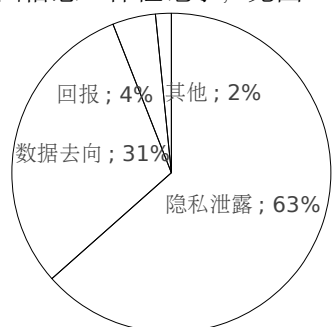


图 1 共享顾虑

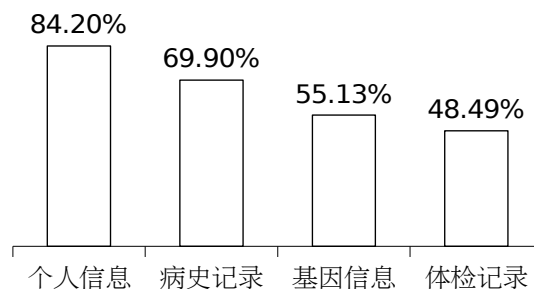


图 2 数据敏感

度认知

3.2 单因素分析

采用卡方检验对居民的共享意愿进行单因素分析，结果显示：学历、数据脱敏需求、签署同意书需求、法律保护需求、回报需求、利于科学研究、利于健康管理是影响健康医疗数据共享意愿的相关因素（ $P<0.05$ ），见表 2。性别、年龄、顾虑对数据共享意愿无统计学意义（ $P>0.05$ ）。

表 2 居民健康医疗数据共享意愿单因素分析

变量	类别	愿意		不愿意		χ^2	P 值
		频数	百分比 (%)	频数	百分比 (%)		
性别	男	195	83.33	39	16.67	0.04	0.84
	女	501	84.20	94	15.80		
年龄	18 岁以下	37	78.72	10	21.28	4.0396	0.25
	18~40 岁	618	84.89	110	15.11		
	41~60 岁	23	76.67	7	23.33		
	60 岁以上	18	75.00	6	25.00		
学历	本科以下	110	77.46	32	22.54	4.795	0.02
	本科及以上	586	85.30	101	14.70		
顾虑	有	631	83.58	124	16.42	0.619	0.43
	无	65	87.84	9	12.16		
数据脱敏需求	是	639	85.20	111	14.80	9.79	0.007
	否	31	75.61	10	24.39		
	不清楚	36	75.00	12	25.00		
签署同意书需求	是	659	84.70	119	15.30	15.751	$P<0.001$
	否	25	83.33	5	16.67		
	不清楚	9	50.00	9	50.00		
法律保护需求	是	675	84.69	122	15.31	9.062	0.01
	否	14	70.00	6	30.00		
	不清楚	7	58.33	5	41.67		
回报需求	是	355	84.93	63	15.07	15	$P<0.001$
	否	160	90.91	16	9.09		
	不清楚	181	77.02	54	22.98		
利于科学研究	同意	658	87.04	98	12.96	65.396	$P<0.001$
	不同意	5	33.33	10	66.67		
	不清楚	33	56.90	25	43.10		
利于健康管理	同意	623	88.12	84	11.88	65.384	$P<0.001$
	不同意	15	48.39	16	51.61		
	不清楚	58	63.74	33	36.26		

3.3 Logistic 回归分析

以居民健康医疗数据共享意愿为因变量（愿意=1，不愿意=0），将卡方检验有差异的变量（学历、数据脱敏需求、签署同意书需求、法律保护需求、回报需求、利于科学研究、利于健康管理）作为自变量，采用二分类 Logistics 向前逐步回归构建回归方程，进入水平 $p<0.05$ 。分析结果显示，签署同意书需求、数据脱敏需求、利于科学研究、利于健康管理四个因素对数据共享意愿有统计学意义。其中，对于健康医疗数据共享要求签署同意书和进行数据脱敏的人群共享意愿更高，认为健康医疗数据共享有利于科学研究和健康管理的人群更愿意实现健康医疗数据共享，见表 3。

表 3 居民健康医疗数据共享意愿多因素 Logistic 分析

影响因素	β	p	OR	95%CI	
				下限	上限
签署同意书需求	-0.934	0.028	0.393	0.16	0.859
数据脱敏需求	-0.385	0.046	0.681	0.471	1.006
利于科学研究	-0.590	$P<0.001$	0.554	0.402	0.768
利于健康管理	-0.572	$P<0.001$	0.564	0.433	0.742
常量	5.416	$P<0.001$	2.93	-	-

4 讨论与建议

4.1 出台健康医疗大数据法，细化数据保护制度

2021年8月20日通过的《中华人民共和国个人信息保护法》，2021年11月1日起正式施行^[5]。本法明确规定，任何组织、个人不得非法收集、使用、加工、传输他人个人信息，不得非法买卖、提供或者公开他人个人信息。在第二章个人信息处理规则中，除规定情形以外，处理个人信息应取得个人同意，并有权撤回同意。处理敏感个人信息应当取得个人的单独同意，而只有在具有特定的目的和充分的必要性，并采取严格保护措施的情形下，才可处理敏感信息。在第四章个人的权利中规定个人对其个人信息的处理享有知情权、决定权，有权限制或者拒绝他人对其个人信息进行处理。

《个人信息保护法》的出台，为数据共享提供了法律保障。但我国在健康医疗数据共享的管理办法和实施细则尚不健全，今后应紧密结合医疗健康行业的发展特点，出台具有中国特色的、针对性的、系统性的健康医疗大数据法，细化个人健康医疗数据保护制度，解决数据的界定、责权、加工、共享、交易等问题，真正实现数据的共享，充分发挥健康医疗大数据的利用价值。

4.2 发展隐私安全计算技术，保障数据安全

随着人们对隐私保护意识的加强，如对个人信息保护过于敏感，将会影响个人信息的合法有效利用。利用隐私安全计算技术如数据脱敏、匿名化、差分隐私、同态加密和联邦学习等^[6]，实现在共享个人健康医疗数据的同时，不会造成个人隐私泄露和信息安全隐患，打消人们对于数据共享的顾虑，建立敏感信息安全的信心，将会保障数据共享顺利开展。对于匿名处理医疗数据的立法，可以借鉴日本的“医疗大数据法”，制定基本方针、认定匿名加工医疗信息的制作人、以及对匿名加工医疗信息的交易等进行规定，促进健康医疗相关的尖端研究开发及新产业创造，进而推动全面大健康的实现^[7]。

4.3 做好健康医疗数据宣传，激发共享积极性

为了健康医疗事业的快速发展，作为健康医疗数据的产生者之一有义务提供健康医疗数据用于科学研究和临床实践。目前广大居民对于数据共享知之甚少，有关部门应做好健康医疗数据知识和相关法律条例的普及，宣传数据共享在科研、临床实践以及健康管理方面的作用，打消隐私泄露顾虑，建立数据共享的信心，激发数据共享的积极性。

5 结语

本次研究仅针对居民健康医疗数据共享意愿进行调查，分析了影响共享意愿的关键因素，并为推动居民对健康医疗大数据的积极共享提出了相关建议。健康医疗大数据的发展面临着平衡数据应用与隐私安全保障的需求，随着法律法规逐步完善、隐私保护技术的提升，健康医疗大数据将得以有效利用，支持医学科学研究，促进医疗健康事业快速发展。

参考文献

- 1[]关于印发国家健康医疗大数据标准、安全和服务管理办法(试行)的通知[J]. 中华人民共和国国家卫生健康委员会公报. 2018(07):7-11.
- 2[] 孟群,毕丹,张一鸣,尹新,等. 健康医疗大数据的发展现状与应用模式研究[J]. 中国卫生信息管理杂志,2016,13(6):547-552.
- 3[]汪冬,秦利,魏洪河,邵泽国. 健康医疗大数据发展现状与应用[J]. 电子技术与软件工程,2018,(11):241-242.
- 4[] 石晶金,于广军. 健康医疗大数据共享关键问题及对策[J]. 中国卫生资源,2021,24(03):223-227+237.
- 5[] 新华社. 中华人民共和国个人信息保护法[EB/OL],[2021-10-25]
http://www.gov.cn/xinwen/2021-08/20/content_5632486.htm
- 6[] 邢丹,徐琦,姚俊明. 边缘计算环境下基于区块链和联邦学习的医疗健康数据共享模型[J]. 医学信息学杂志,2021,42(02):33-37
- 7[] 廖子锐,田雪晴,刘远立. 日本医疗大数据法对我国健康医疗大数据发展应用的启示[J]. 中国数字医学,2021,16(07):88-93.