

卫生健康数据开发利用的国内外经验与启示

汪 茹^{1*} 冯 骏² 蒋璐伊^{1,3}

1. 上海市卫生和健康发展研究中心(上海市医学科学技术情报研究所) 上海 200031
2. 上海市卫生健康委员会 上海 200125
3. 复旦大学上海市重大传染病和生物安全研究院 上海 200032

【摘 要】数据已成为我国第五大生产要素,为加快公共数据资源开发利用,我国已形成“1+3”的政策体系框架,然而在数据授权定价、授权模式等方面仍存在较多探索空间,尤其卫生健康数据兼具个人隐私与公共数据属性,在安全保护与开发效益之间存在许多因素需要权衡。英国、美国及我国香港地区在卫生健康数据的开发利用方面有一定的先进经验,本文系统研究梳理了这些国家和地区在卫生健康数据开放平台建设、组织架构、运营机制等方面的制度设计,结合我国实际,在卫生健康数据平台建设、数据资源登记、伦理审查与安全管理以及定价模式等方面提出具体意见,以期相关政策制定提供参考。

【关键词】卫生健康数据; 开发利用; 授权运营

中图分类号:R197 文献标识码:A doi:10.3969/j.issn.1674-2982.2025.10.009

Domestic and international experience and implications on development and utilization of health data

WANG Ru¹, FENG Jun², JIANG Lu-yi^{1,3}

1. Shanghai Health Development Research Center (Shanghai Medical Information Center), Shanghai 200031, China
2. Shanghai Municipal Health Commission, Shanghai 200125, China
3. Shanghai Institute of Infectious Disease and Biosecurity, Fudan University, Shanghai 200032, China

【Abstract】Data has become the fifth major factor of production in China. To accelerate the development and utilization of public data resources, China has established a "1+3" policy system framework. However, there remains considerable room for exploration in aspects such as data authorization pricing and authorization models. In particular, health data possesses both personal privacy and public data attributes, requiring multi-factor trade-offs between security protection and development benefits. The United Kingdom, the United States, and China's Hong Kong Special Administrative Region have accumulated certain prior experiences in the development and utilization of health data. This article systematically studied and sorted out the institutional design of the construction, organizational structure, and operation mechanism of health data open platforms in these three countries/regions. Combining with China's domestic realities, specific suggestions were put forward regarding the construction of health data platforms, data resource registration, ethical review and security management, as well as pricing mode, aiming to provide a reference basis for the formulation of relevant policies and systems.

【Key words】Health data; Development and utilization; Authorized operation

1 研究背景

信息技术革命和人类经济社会活动的交汇融

合,引发了全球数据爆发式增长,数据已成为继土地、劳动力、资本、技术之后的第五大生产要素^[1],是发展新质生产力的重要基础^[2]。卫生健康数据是指

* 基金项目:上海市卫生健康委员会2024年卫生健康政策研究课题(2024HP48);上海市卫生健康委员会2025年卫生健康政策研究课题(2025HP31)

作者简介:汪茹(1996年—),女,硕士,研究实习员,主要研究方向为数字健康、卫生政策研究、卫生技术评估等。E-mail:wangru@shdrc.org
通讯作者:蒋璐伊。E-mail:elenasweet@163.com

在疾病防治、健康管理、医学相关教学研究、医疗管理等过程中产生的数据^[3],被视为国家重要的基础性战略资源^[4]。随着传统医药卫生行业与人工智能等新兴技术的加速融合,卫生健康数据呈现爆发式增长态势,逐步发展成为新业态与新经济增长点。^[4]

促进数据开发利用,发掘数字要素价值,构建数字资产,已逐步成为全球共识。^[5]为推动和规范数据的利用与发展,我国加快了政策制定与行业布局。2022年12月,中共中央、国务院发布《关于构建数据基础制度更好发挥数据要素作用的意见》,要建立公共数据、企业数据、个人数据的分类分级确权授权制度,建立数据资源持有权、数据加工使用权、数据产品经营权等分置的产权运行机制,推进非公共数据按市场化方式“共同使用、共享收益”的新模式,加强公共数据汇聚共享和开放开发,强化统筹授权使用和管理,推进互联互通,打破“数据孤岛”。^[6]

目前,在非公共数据方面,国内已上线60多家数据交易所、中心、平台,其中有多家开设生物医药专区,如上海数据交易所甄选260余个生物医药相关的数据集、数据服务、数据应用挂网展示,包括上市药品专利、医保药品信息、疾病分析、临床试验测试和行业分析等数据,以及医药器械、智慧诊疗和健康管理等典型应用场景。^[7-8]在公共数据方面,2024年9月,中共中央办公厅、国务院办公厅发布《关于加快公共数据资源开发利用的意见》^[9],首次在中央层面对公共数据资源开发利用进行系统部署。国家发展改革委、国家数据局随后陆续公布了《公共数据资源登记管理暂行办法》(发改数据规〔2025〕26号)^[10]、《公共数据资源授权运营实施规范(试行)》(发改数据规〔2025〕27号)^[11]以及《关于建立公共数据资源授权运营价格形成机制的通知》(发改价格〔2025〕65号)^[12],自此,我国公共数据开发利用形成了1个总体纲领加3个配套政策的“1+3”政策体系。地方政策也逐步跟进,如上海市在《上海市公共数据资源授权运营管理办法》(沪府办发〔2025〕15号)中提到公共数据采用统一授权模式^[13],北京市在《关于加快北京市公共数据资源开发利用的实施意见》中表示要积极推动公共数据分领域授权、依场景授权,逐步探索整体授权运营。^[14]

总体来看,我国公共数据开发利用工作已完成顶层框架设计,但在落地实施方面还存在各地个性化探索实践阶段,同时,数据运营价格等方面还未形成可操作的落地性文件。基于我国公立医疗卫生机

构为主的诊疗现状下,卫生健康数据是我国公共数据的重要组成部分,兼具公民隐私与公共数据属性,如何平衡数据安全与经济效益,是卫生健康数据开发利用工作的关注重点。经文献调研,英国、美国及中国香港地区分别通过构建UK Biobank、Health Data NY、香港数据协作实验室等规范化统一数据开放平台,在数据标准化、平台运营模式和价值转化机制等方面逐步形成了体系化路径,对我国卫生健康数据平台建设及开发利用工作有一定的参考作用,本研究将聚焦上述三个国家和地区的卫生健康数据开放运营实践,提炼可借鉴的管理机制,以期为相关政策制定提供参考。

2 研究方法

2.1 检索策略

本研究基于网络搜索引擎进行全球范围内卫生健康数据开发利用情况面上检索,通过基础情况梳理遴选几个典型模式进行专题深入分析。首先分别以中文关键词“卫生健康数据”“开发利用”“授权运营”“数据平台”和英文关键词“health data”“development and utilization”“authorized operation”“data platform”等进行组合检索,初步将美国Healthdata.gov、美国纽约Health Data NY、美国加利福尼亚州Health and Human Services、英国Data.gov.uk、英国UK Biobank、香港Data.gov.hk、香港数据协作实验室、加拿大Open Data和澳大利亚Data.gov.au几大数据平台纳入信息收集范围。通过对以上数据平台的建设情况和开发利用模式的初步梳理,确定纳入英国UK Biobank、美国纽约Health Data NY和香港数据协作实验室3个典型案例进行分析,并进一步查阅相关平台官网信息及学术文献等资料,3个平台的官方信息来源详见表1。

表1 三个数据平台官网网页信息

国家/地区	平台	来源
英国	UK Biobank	https://www.ukbiobank.ac.uk/
美国	Health Data NY	https://health.data.ny.gov/browse? sortBy=alpha
中国香港	数据协作实验室	https://www3.ha.org.hk/data/

2.2 分析框架

本文以英国、美国、中国香港3个典型国家和地区的卫生健康数据开放平台为研究对象,基于平台

建设、组织架构、运营机制3个维度开展信息收集(图1),通过对比分析梳理3个平台建设的共性与个性特点,结合我国国情提出建议。

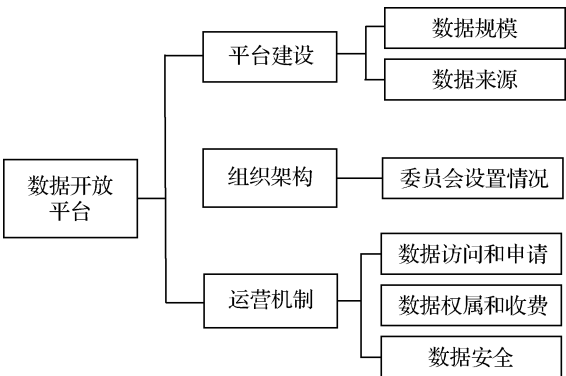


图1 3个数据平台开发利用模式分析框架

3 申请收费模式:英国生物银行(UK Biobank)

3.1 平台建设

英国生物银行(UK Biobank)即英国生物样本库,是一个大型前瞻性队列研究及生物医学数据库,其核心资源的资金支持由英国医学研究委员会、英国卫生部、威康信托基金会、苏格兰行政院和英国西北地区发展局向英国生物样本库有限公司提供。^[15] UK Biobank作为一个大规模的生物医学数据库和研究资源,包含来自英国50万例40~69岁志愿者的去标识化的遗传、生活方式和健康信息以及生物样本,并长期跟踪记录参与者的健康和医疗档案信息。包含精神健康、脑部和心脏磁共振成像、认知测试、心肺功能检测、身体活动记录、血液样本中的基因、生化指标、蛋白质等数据,另外还有关联的疾病诊断、入院登记、癌症或者死亡登记等信息,能够覆盖临床大部分科室。^[16] UK Biobank还将数据按照不同领域整合成了数据集,每一项数据集都有一个唯一的识别ID。该数据库通过提供丰富的生物医学数据,支持全球范围内的研究人员开展健康相关的研究。根据UK Biobank官网统计,目前全球有超过16 000篇同行评审论文使用过该数据。^[17]

3.2 组织架构

UK Biobank设有董事会,董事会成员全面负责UK Biobank的指导与管理监督,首席研究员负责项目设计和各部门管理。UK Biobank作为数据和样本的法定保管主体,运营实体为曼彻斯特大学下设的协调中心。协调中心负责队列全过程运作,包括数

据保管及发布、人员培训、质量保证、财务管理和监督,其下设的6个区域合作中心覆盖了20余所高校,负责志愿者招募、基线样本数据采集、科研方案咨询及项目落地执行和交付。此外,UK Biobank设有四个委员会,分别为准入委员会(Access Committee)、伦理咨询委员会(Ethics Advisory Committee)、信息治理委员会(Information Governance Committee)、审计与风险委员会(Audit and Risk Committee)。其中准入委员会负责数据科学准入的关键决策,并直接向董事会提供科研咨询;伦理咨询委员会独立于其日常管理体系,主要承担三项核心工作:一是就伦理问题向董事会提供咨询并开展审查;二是为相关事务提供政策建议;三是发布公开报告,说明UK Biobank是否符合伦理与治理框架,以及是否维护了参与者和公众的利益;信息治理委员会负责监督和加强UK Biobank在数据收集、存储和发布过程中的策略和运营安全;审计与风险委员会独立于UK Biobank的日常管理,负责审查UK Biobank内部监控系统与风险管理系统是否足够完善和切实有效^[18]。

3.3 运营机制

3.3.1 数据的访问申请流程

第一步:账号注册及审核。研究者需要在UK Biobank官网注册访问管理系统的账号,并进一步完善个人信息,包括个人科研背景和英文履历,完成后会进入“待审核”阶段,由UK Biobank工作人员进行审核,正常情况下需3~5个工作日。第二步:数据使用申请、审查、批准。UK Biobank数据的使用需要依托项目进行申请,需要申请人准备好研究项目的标书,主要包括项目概述、研究人员姓名、项目主要内容等,UK Biobank官网也提供了已批准项目的简明摘要(Lay Summary)介绍可供参考。此时支付预付款250欧元,研究者需先明确申请方向:若仅申请研究数据,按常规流程办理;若需申请生物样本数据,则该类申请需通过更严格的科学审查与伦理审查。申请提交后,UK Biobank的工作人员会对申请书进行严格审核,审核时间大约4周。第三步:提交数据使用协议(Material Transfer Agreement, MTA)并支付数据使用费,目前UK Biobank数据使用费分为3档。申请获批后,UK Biobank的工作人员会以电子邮件形式通知,申请人需要根据要求线上填写MTA并付费,审核时间大约4周。第四步:获取数据。UK Biobank的数据将在云计算平台中提供,某些数据也

可以下载供已获批准的研究人员在特定项目中使用。^[19]

3.3.2 数据权属和收费

目前,UK Biobank 已经形成比较成熟的运营和收费机制。根据《英国生物样本库伦理和监管框架》,UK Biobank 拥有数据产权,参与建设者拥有数据使用的知情同意权,研究者的研究成果与 UK Biobank 共有。对于参与建设者,UK Biobank 提供长达 9 个月的数据免费访问期。9 个月之后,参与建设者则需要和其他研究人员一样进行付费申请,一次申请可获得 1~3 年的数据访问期限(表 2)。在收费方面,UK Biobank 根据数据的不同类型设置了不同档位的收费价格。第一档为核心数据,如问卷调查、身体检查、关联健康数据、健康结果表型等,可根据数据的具体内容收费 3 000、6 000 或 9 000 英镑;第二档为化验数据,如生化和血液分析、基因数据等,可根据数据的具体内容收费 6 000 或 9 000 英镑;第三档为超大数据,如影像数据、全基因组序列、其他大规模化验数据、全外显子组序列等,收费 9 000 英镑。此外,若有其他的研究机构加入,前三年需额外支付 1 000 英镑;若申请延期,每年需支付 500 英镑延期费。此外,学生及来自低收入国家的研究人员可享受访问费减免与财政支持,具体标准为前三年仅需支付 500 英镑,若申请延期,每年仅需支付 175 英镑。^[20]

表 2 不同对象的数据使用标准

开放贡献	参与建设者	其他研究者
对象	提供资金支持的企业、基金会及其研究人员	全球学术、商业和慈善组织的研究人员
期限	免费访问期(9个月)	3年(可申请延长1~3年)
范围	所有数据	数据分类(3类):核心数据、化验数据、超大数据
费用	免费	收费(3档,学生和低收入国家研究人员可申请部分减免)

3.3.3 数据安全

UK Biobank 的数据审核和隐私保护遵循《英国生物样本库伦理和监管框架》。具体包括(1)数据匿名化:UK Biobank 将数据和样本以可逆匿名化的形式安全存储,参与者在评估访问之后,身份识别信息将转移到 UK Biobank 中央系统进行数据脱敏,所有参与者的个人身份信息与数据分开,并使用没有任何含义的代码进行代替,所有身份识别信息将由 UK

Biobank 集中保存在一个访问受限的数据库中,该数据库由 UK Biobank 的高级职员控制,只有少数人能够获得“密钥”,研究人员只能获得匿名化的样本数据。(2)研究人员审查:研究人员必须经过 UK Biobank 团队的审查,确保他们来自可信任的研究组织,并且其研究项目符合公共利益。(3)伦理审查:在获得访问权限之前,所有研究项目必须经过 UK Biobank 的审查,以确保其获得参与者的同意并符合伦理框架。(4)签署法律协议:一旦研究申请被批准,研究组织必须签署法律协议才能访问数据。(5)禁止识别参与者身份:研究人员禁止利用 UK Biobank 的任何数据(包括遗传数据)来识别参与者的身份。^[21]

4 网站开放模式:纽约健康数据(Health Data NY)

4.1 平台建设

2009 年,奥巴马政府发布的《开放数据指令》要求政府部门设立开放数据专属网页,促进公共领域对话。为响应该政策,纽约州发起了“OPEN NY”倡议,并于 2013 年 3 月 11 日推出纽约州开放数据网站 open. ny. gov。其中,Health Data NY 是“OPEN NY”项目中为提升卫生健康数据透明度而建立的一个纽约州级的健康数据开放平台,由纽约州卫生部门承担主要的建设任务。该平台包含 537 个数据条目,包括数据集、图表和地图等。这些资料主要来自于纽约州卫生部门的调查、疾病报告和重要记录等,涉及环境健康、医疗机构与服务、社区健康与慢性病、医疗质量安全与成本、出生死亡记录等健康相关数据。^[22]

4.2 组织架构

Health Data NY 平台以数据工作组的方式开展工作。数据工作组由来自纽约州信息技术服务办公室(Office of Information Technology Services, ITS)的代表及各州代表组成。其中,ITS 负责管理 Health Data NY 平台,并维护 open. ny. gov 和 health. data. ny. gov 网站,为涉及的机构和州提供指导和技术支持。ITS 内部会指定一位首席数据官,负责协调开放数据网站的具体实施工作。数据工作组会定期发布数据指南,指导提供数据的各州机构。提供数据的州机构会按照指南要求创建一份其可公布的数据目录,并向 ITS 和首席数据官提出一份公开计划表,以公开其可公布的州数据,并定时更新数据目录。此外,提供数据的各州机构会指派一位数据协调员作为 ITS 和

所在机构之间的联络人,负责获得所在机构对数据发布、更改和删除的批准,并与 ITS 合作发布和更新数据,保障各州能够高效的跨部门合作(图 2)。

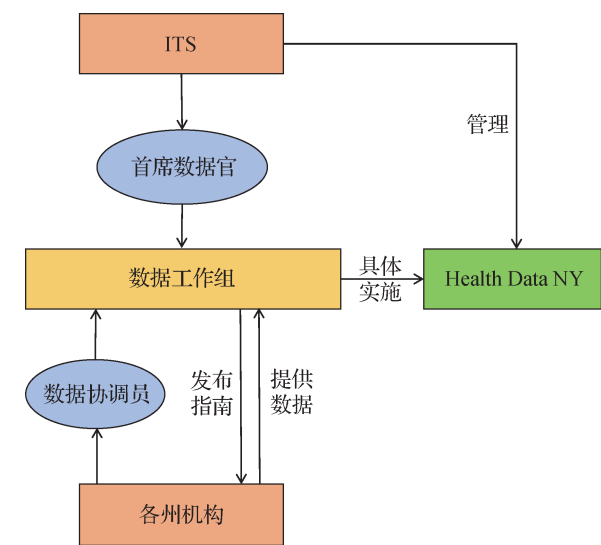


图 2 Health Data NY 管理和组织架构

4.3 运营机制

4.3.1 数据的访问和申请

Health Data NY 平台提供了一个开放的应用程序接口 (Application Program Interface, API), 无需在线申请, 开发人员可通过 API 对原始数据进行直接访问, 并动态查询数据目录中的数据, 且以 JSON、XML、CSV 等文件格式获取数据。

4.3.2 数据权属和收费

目前, 纽约健康数据网站提供免费数据下载服务。在数据权属方面, Health Data NY 平台数据属于纽约州卫生部, 并保留对所访问数据集的版本控制。但需注意, 研究人员一旦下载数据, 纽约州卫生部将不对此数据作出任何担保或质量保证, 也不对该网站检索数据所开展的任何分析承担担保责任。

4.3.3 数据安全

ITS 设有首席安全官, 各州机构设有安全和隐私官员, 负责确保各州和机构发布数据时的数据安全。Health Data NY 平台要求, 纽约州及各机构主体收集的数据, 必须保证是在遵守所有适用法律、规则、法规、条例、决议、政策或其他与州数据相关的限制、要求的情况下, 可以向公众提供的数据, 并严格遵守隐私安全要求、知识产权和《信息自由法》的规定。

5 实验室模式: 香港数据协作实验室 (Data Collaboration Lab)

5.1 平台建设

香港数据协作实验室由香港医院管理局 (Hospital Authority, HA) 成立, 该实验室是一个受控制的物理环境, 数据只能在实验室内处理分析, 相关数据一律不能带离实验室或制作复本。数据协作实验室拥有自 1997 年起的香港相关医疗健康数据, 具体涵盖四大类: 一是人口特征数据; 二是就医记录, 包括医院及诊所就诊情况; 三是诊疗数据, 包含临床诊断、手术及其他医疗程序; 四是辅助诊疗数据, 涉及药物信息、化验结果以及放射检查的图像与报告。目前, 数据协作实验室通过数据协作项目和自助数据平台进行数据开放。数据协作项目主要面向香港本地的科研和学术机构, 实验室会根据项目要求提供研究人员所需的数据、分析工具及技术平台, 以便进行更深入的数据分析。同时, 医管局设有医疗信息专业人员作为数据协作者支持或参与研究项目。目前, 通过数据合作项目发表的论文数量已达 42 篇。此外, 数据协作实验室设有自助数据平台, 包含约 20 万名匿名化患者的临床数据集, 样本数据集由 2007—2017 年从按年龄及性别划分的患者中随机抽取的结构化数据构成。^[23-24] 自助数据平台主要帮助研究人员进行探索研究, 开展研究设计和预实验, 一是测试研究项目所需的数据是否足够, 以确定纳入研究的样本数量, 二是探索所需的目标风险因素是否包含在数据库内, 以便对项目计划进行调整。此外, 自助数据平台还提供数据分析工具供研究人员在线开展数据处理。自助数据平台的用户群体, 除了香港本地研究机构的科研人员外, 还进一步拓展到了创新科技领域从事相关工作的专业人员或从业者。2024 年 3 月 28 日, 医管局与香港科技园公司 (Hong Kong Science and Technology Parks Corporation, HKSTP) 合作, 于香港科技园内设置科技园—医管局数据实验室, 供园区内指定的生物科技及医疗保健公司使用, 以推动官产学研协作。^[24]

5.2 组织架构

HA 设有数据实验室评审委员会, 对项目的数据需求、合理性及伦理合规性进行评审, 以确保数据使用严格遵循安全和隐私保护原则, 并符合临床研究伦理标准。此外, HA 的数据实验室办公室负责安排

实验室的使用,同时,香港科技园设置临床研究伦理委员会和数据管理委员会,确保企业使用数据时均遵循上述原则和标准。

5.3 运营机制

5.3.1 数据的访问和申请

数据协作实验室的数据申请根据申请人不同,流程上有所区分。一是针对数据协作项目的申请者。第一步:数据使用申请。已与医管局签订合作协议的大学研究人员可在线填写申请表,除了项目的详细信息外,还需添加数据需求,包括所需的数据结构并填写具体的纳入和排除标准,如样本的采集时间、患者的年龄、纳入的疾病等信息。根据选择所需的数据目录和对数据集的具体要求,不在数据目录中的数据需额外填写。同时,申请者还需要提交伦理审批结果,伦理审批由医管局认可的伦理事务委员会出具,包括医管局辖下的医院研究伦理委员会及本地大学的伦理事务委员会。另外,申请者还需提供显示该研究项目已获得资助的机构支持证明。第二步:审查、批准评审。成功递交申请的研究

人员将会收到附有申请编号的确认邮件,评审委员会就申请作出评审。通过评审的项目,办公室会邮件通知申请人,并告知研究人员使用数据实验室的期限。需要注意的是,申请人最多可以申请5天使用期,若未能于该期限内完成数据分析,研究人员需要重新申请使用自助数据平台,并向办公室申请延长研究工作的保留期。第三步:签订合作协议及承诺书。研究人员及所属机构需与医管局签订合作协议及承诺书,办公室在收到同意书后会准备研究项目所需的数据和分析工具,并在完成准备工作后通知研究人员有关使用数据实验室的安排。第四步:研究发布。在符合协议及承诺书内确定的数据保密要求下,研究人员可于最后使用自助服务日起一个月內申请输出数据和脚本,亦可发布研究。^[25]二是针对使用自助平台的研究人员及科学院园区内的企业。首先需填写申请表格以及保密承诺书,并提交给大学的联络点或园区的办公室。若使用自助平台的数据进行研究成果发布,则需进行额外的临床研究伦理委员会审批^[26],其它流程与数据协作项目申请流程类似(图3)。

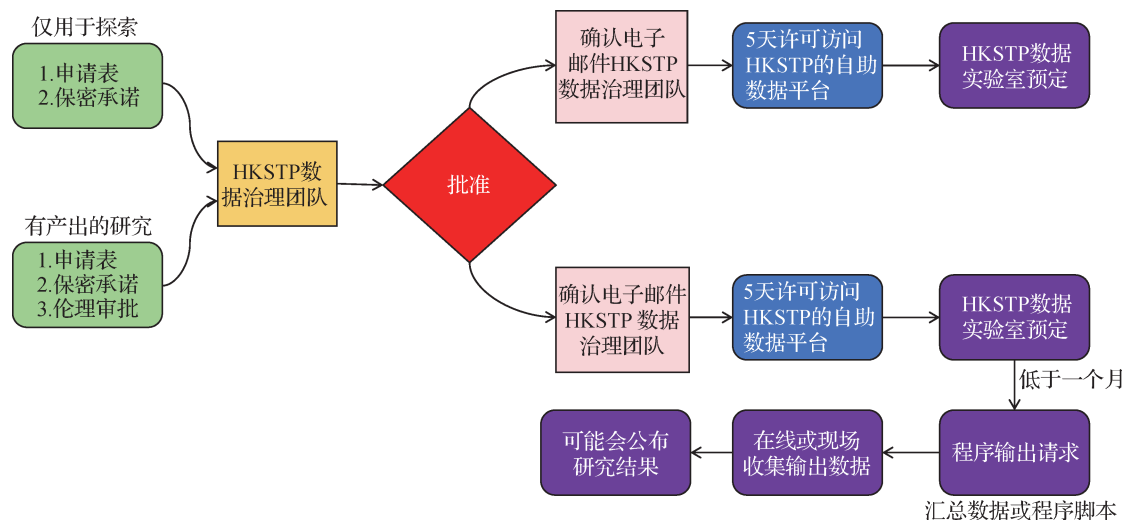


图3 HA数据实验室数据使用申请流程图^[27]

5.3.2 数据权属和收费

目前,香港数据协作实验室的数据申请服务有两种收费选项,一是快速申请,涵盖常用的数据项,处理申请的费用通常为15 000美元及以上,具体取决于数据选择和规格;二是定制化申请,涵盖申请表覆盖范围外的数据项,处理申请的费用通常为60 000美元及以上,具体取决于数据的复杂性。数据协作项目仅对已与医管局签订合作协议的大学研究人员

开放,自助数据平台对以上研究人员或香港科技园区内的指定公司开放。

5.3.3 数据安全

科技园区公司采取将数据留存在物理空间的严格安全措施保护数据安全和隐私。具措施包括安排当值工作人员、安装监控系统并设置专用房间,严格控制人员进出。此外,还包括禁止连接互联网,所有数据严禁下载、保存或打印,确保所有数据只会保留

在医管局,其余机构或公司无法下载、复制或存档,只能以样本数据作研究。

6 比较分析

综合分析英国 UK Biobank 申请收费模式、美国 Health Data NY 免费开放模式和香港数据协作实验室模式,发现三种模式具有共性,又各有特点。在组织架构上,UK Biobank 和香港数据协作实验室均设有委员会,其中,UK Biobank 的伦理咨询委员会与香港数据协作实验室的临床研究伦理委员会功能类似,均负责项目中涉及的医学伦理审核;两者分别设置准入委员会和数据管理委员会负责项目内容审查和数据安全。而 Health Data NY 则以数据工作组模式开展工作,由各州派遣一位数据联络员与 ITS 代表共同组成,相比而言,Health Data NY 没有专人负责伦理审核。在运营和收费模式上,UK Biobank 按照数据量、数据规格、使用时长和使用对象不同设置分级分类收费标准,目前其运营模式已较为成熟,香港数据协作实验室根据数据规格和复杂程度有两种

收费选项,而 Health Data NY 免费提供数据查询和下载服务。在使用期限上,UK Biobank 为 3 年,延期需要额外付费;香港数据协作实验室为 5 天,逾期需要提出申请;Health Data NY 则没有限制。在数据规模上,三个平台的数据量、数据性质和涵盖范围不同。UK Biobank 是一个包含了 50 万例志愿者去标识化信息的前瞻性大型队列研究及生物医学数据库,Health Data NY 和香港数据协作实验室收集的是回顾性数据,相比而言,UK Biobank 的数据质量最高。此外,UK Biobank 和香港数据协作实验室仅涵盖医学相关的数据,Health Data NY 还涵盖了医疗保险数据、心理健康、环境健康等数据。在数据安全上,UK Biobank 和 Health Data NY 分别对卫生健康数据进行了匿名化和脱敏处理,香港数据协作实验室则是通过物理隔离的方式保证数据安全。在结果产出上,利用 UK Biobank 数据库进行研究论文产出 10 000 多篇,香港数据协作实验室 42 篇,Health Data NY 暂无学术产出报道(表 3)。

表 3 各国家或地区卫生健康数据平台基本情况对比

国家或地区	平台	面向对象	组织架构	运营模式	使用期限	数据量	数据来源	数据类型	数据质量	数据安全
英国	UK Biobank	全球范围内的研究人员	设有董事会、伦理咨询委员会、准入委员会、信息治理委员会、审计与风险委员会	分级分类收费	3 年	50 万例	英国 40~69 岁的志愿者	前瞻性队列数据	高	样本信息数据匿名化
美国	Health Data NY	全球范围内的研究人员	数据工作组	免费	永久	533 个条目	纽约州卫生部的调查、疾病报告和重要记录等	历史资料	一般	样本信息数据脱敏
中国香港	数据协作实验室	香港本地的科研和学术界机构、香港科技园区内指定的公司	设有临床研究伦理委员会、数据管理委员会	收费	5 天	20 万例	2007—2017 年按年龄及性别划分的患者中随机抽取的结构化数据	临床结构化数据集	较高	物理空间安全,数据不可带走

7 对我国卫生健康数据开发利用的启示

卫生健康数据是国家重要的基础性战略资源,其开放运营对推动健康中国建设与生物医药产业发展至关重要。^[28]当前,国家级平台设有国家北方健康医疗大数据中心、国家东南健康医疗大数据中心、国家人口与健康科学数据共享平台等。各省市也在加快大数据共享开放平台建设,启动运营的有深圳市政府数据开放平台、苏州市健康医疗数智创新实验

室、湖北省卫生健康委员会建设数据开放平台、杭州市数据开放平台和上海市卫生健康数据平台等,但我国卫生健康数据开发利用仍面临多重挑战:医院“数据孤岛”现象普遍、数据质量参差不齐、审查与数据安全保障机制尚需强化、缺乏成熟的数据定价与收费机制以及企业与研究机构数据利用深度和创新活力不足,均制约着卫生健康数据价值释放。为有效推动我国卫生健康数据开发利用,基于国内外相关经验,本研究提出以下建议。

7.1 加强数据平台建设 with 数据资源登记

统一高质量的数据平台建设是数据开发利用的前提保障和趋势。统一的卫生健康数据平台建设有利于破除数据孤岛、数据标准不一以及质量参差不齐等问题。基于平台建设,可对历史归集的数据进行系统化治理,包括数据标准化、结构化处理以及跨平台兼容性建设,以增强数据的可用性和研究价值。借鉴 UK Biobank 依靠高质量前瞻性队列数据驱动研究产出经验,我国可在提升回顾性数据质量的基础上,探索实施前瞻性的数据采集工作,通过促进高质量学术成果产出及产业应用转化,加强平台的使用价值和品质提升。同时,加强数据资源登记以提高数据资源的透明度与可访问性。参考三个国家和地区的经验,建立清晰的数据资源目录,可提升大众对数据资源的认知度和使用度。如 UK Biobank 和香港数据协作实验室均在申请时面向申请人提供结构化数据目录,Health Data NY 通过开放数据平台向公众提供政府健康信息,提高社会对健康数据的认知,可借鉴探索建立国家级、区域级数据开放门户,向公众提供明晰的数据资源目录与健康信息,提升社会认知度并促进合理利用。

7.2 强化伦理审查与数据安全治理

卫生健康数据的安全合规流通应用,需要强化伦理审查和数据安全管理。UK Biobank 与香港数据协作实验室均设有独立的伦理委员会和数据科学委员会,分别负责项目的伦理审查和科学合理性评估。结合国内外的经验,提示应加强伦理与技术双审核管理。在伦理要求方面,我国已有明确政策要求,根据国家卫生健康委、教育部、科技部、国家中医药局《关于印发涉及人的生命科学和医学研究伦理审查办法的通知》(国卫科教发〔2023〕4 号),以人为受试者或者使用人(统称研究参与者)的生物样本、信息数据(包括健康记录、行为等)开展的研究活动应当进行医学伦理审查。^[29]在技术安全管理方面,三个国家和地区在数据安全管理方面采取了多层次的技术和制度保障措施。UK Biobank 设立了信息治理委员会负责保障数据安全,并建立了严格的数据安全保护机制;Health Data NY 在治理上设置了首席安全官和地方的安全和隐私官员,在数据上传之前就确保其符合法律和隐私要求;香港数据实验室采用物理隔离技术以确保数据安全。相关经验提示我国在构建卫生健康数据安全体系时,可结合本土需求,充分

利用云计算、区块链等新技术,加强数据加密、访问控制及可追溯性管理,以确保数据存储与传输的安全性。同时,在人员治理组织架构上,可以加强数据战略与安全管理专员设置。

7.3 探索数据开发利用的多样化定价收费机制

卫生健康数据的开发利用遵循我国公共数据资源授权运营价格形成机制,对于公共治理、公益事业的,可免费提供;用于产业发展、行业发展的,可收取公共数据运营服务费,该费用实行政府指导价管理。合理的价格收费有益于卫生健康数据的可持续性运营。对比 UK Biobank 和 Health Data NY,UK Biobank 实行数据收费模式很大程度上确保了其高质量数据的建设和维护,相比之下,Health Data NY 虽为免费开放模式,但数据品质不高。中国在顶层设计上已根据数据目的区分了免费和收费模式,在细化落地方面,建议可综合实行以下几种模式:一是有条件的共享开放模式。加强政府部门间及行业内部的数据共享流通利用,面向社会提供一些免费的基础性数据,如统计性的卫生健康指标数据、疾病流行趋势等,促进公益性研究和政策规划等制定。二是数据风险分级收费模式。借鉴 UK Biobank 收费模式经验,根据数据的隐私敏感度、数据规模及用途制定分级收费标准。如划分共享数据、开放数据、授权运营数据、限制使用数据等,结合卫生健康数据分类分级标准及数据价值进行综合风险分级收费。三是用户分类收费模式。参考 UK Biobank 对低收入国家及学生的优惠政策,以及香港数据协作实验室针对性的开放权限设置,可探索针对不同类别的研究人员,如科研机构、产业、政策制定者,设立相应的访问权限和访问周期。同时,结合国内市场需求,制定科学合理的分类或优惠定价体系,提升数据的可及性与利用率。四是增值服务收费。针对市场应用需求提供个性化增值服务,如数据挖掘、大模型应用等,可探索市场化高价值收费模式,促进产业发展。

作者贡献:汪茹负责研究设计、定量与定性资料收集与分析、论文撰写与修改;冯骏负责理论指导、研究设计、论文审阅;蒋璐伊负责理论指导、研究设计、论文撰写、论文审阅与修订。

作者声明本文无实际或潜在的利益冲突。

参 考 文 献

[1] 中共中央国务院. 关于构建更加完善的要素市场化配置

- 体制机制的意见[EB/OL]. (2020-03-30)[2025-08-28]. https://www.gov.cn/zhengce/2020-04/09/content_5500622.htm
- [2] 国家数据局. 摸清数据资源底数,释放数据要素价值潜能[EB/OL]. (2024-05-30)[2025-08-28]. https://www.nda.gov.cn/sjj/zwgk/zjjd/0830/20240830190958798716841_pc.html
- [3] 上海市市场监督管理局. 卫生健康数据分类分级要求: DB31/T 1545-2025[S]. 2025.
- [4] 国务院办公厅. 关于促进和规范健康医疗大数据应用发展的指导意见(国办发[2016]47号)[EB/OL]. (2016-06-24)[2025-08-28]. https://www.gov.cn/zhengce/content/2016-06/24/content_5085091.htm
- [5] KOSTKOVA P, BREWER H, DE LUSIGNAN S, et al. Who Owns the Data? Open Data for Healthcare[J]. Front Public Health, 2016, 1(4): 7-12.
- [6] 中共中央, 国务院. 关于构建数据基础制度更好发挥数据要素作用的意见[EB/OL]. (2022-12-02)[2025-07-14]. https://www.gov.cn/gongbao/content/2023/content_5736707.htm
- [7] 60家全国数据交易所/中心/平台名单整理(上海数据交易所正式开设生物医药专区)[EB/OL]. (2024-08-31)[2025-08-28]. <https://mp.weixin.qq.com/s/4qYcdIzXDS4V6Qk1W2UFww>
- [8] 上海市数据交易所. 数据产品交易挂牌流程全指引[EB/OL]. [2025-08-28]. <https://nidts.chinadep.com/ep-hall?search>
- [9] 中共中央办公厅, 国务院办公厅. 关于加快公共数据资源开发利用的意见[EB/OL]. (2024-09-21)[2025-08-28]. https://www.gov.cn/zhengce/202410/content_6978911.htm
- [10] 国家发展改革委, 国家数据局. 关于印发《公共数据资源登记管理暂行办法》的通知(发改数据规[2025]26号)[EB/OL]. (2025-01-08)[2025-08-28]. https://www.gov.cn/gongbao/2025/issue_11886/202502/content_7007130.html
- [11] 国家发展改革委, 国家数据局. 关于印发《公共数据资源授权运营实施规范(试行)》的通知(发改数据规[2025]27号)[EB/OL]. (2025-01-08)[2025-08-28]. https://www.gov.cn/gongbao/2025/issue_11886/202502/content_7007129.html
- [12] 国家发展改革委, 国家数据局. 关于建立公共数据资源授权运营价格形成机制的通知(发改价格[2025]65号)[EB/OL]. (2025-01-16)[2025-08-28]. <https://zfxxgk.ndrc.gov.cn/web/iteminfo.jsp?id=20477>
- [13] 上海市人民政府办公厅. 关于印发《上海市公共数据资源授权运营管理办法》的通知(沪府办发[2025]15号)[EB/OL]. (2025-07-15)[2025-08-28]. <https://www.shanghai.gov.cn/nw12344/20250801/e03bd8f1f16b462ba100d3f3d431c874.html>
- [14] 中共北京市委办公厅, 北京市人民政府办公厅. 关于加
- 快北京市公共数据资源开发利用的实施意见[EB/OL]. (2025-05-06)[2025-08-28]. https://www.beijing.gov.cn/zhengce/zhengcefagui/202508/t20250812_4171531.html
- [15] PALMER L J. UK Biobank: bank on it[J]. Lancet, 2007, 369(9578): 1980-1982.
- [16] UK biokank[EB/OL]. [2025-08-28]. <https://www.ukbiobank.ac.uk/about-our-data/biological-samples/>
- [17] UK biokank[EB/OL]. [2025-08-28]. <https://www.ukbiobank.ac.uk/use-our-data/apply-for-access/>
- [18] UK biokank[EB/OL]. [2025-08-28]. <https://www.ukbiobank.ac.uk/about-us/people-and-governance/>
- [19] UK biokank[EB/OL]. [2025-08-28]. <https://community.ukbiobank.ac.uk/hc/en-gb/categories/14494598931229-Use-our-data>
- [20] UK biokank[EB/OL]. [2025-08-28]. <https://www.ukbiobank.ac.uk/use-our-data/fees/>
- [21] UK biokank[EB/OL]. [2025-08-28]. <https://www.ukbiobank.ac.uk/about-our-data/protecting-the-data/>
- [22] Health Data NY[EB/OL]. [2025-08-28]. <https://health.data.ny.gov/browse?sortBy=alpha>
- [23] Self-service Data Platform[EB/OL]. [2025-06-08]. <https://www3.ha.org.hk/data/DCL/SelfService>
- [24] 医院管理局与科技园正式开放数据平台支援科学园创科企业进行科研[EB/OL]. (2024-03-28)[2025-06-08]. <https://www.hkstp.org/zh-cn/park-life/news-and-events/news/hospital-authority-and-hkstp-formally-launch-data-platform-to-support-scientific-research>
- [25] Hospital Authority/Data Sharing Protal[EB/OL]. [2025-06-08]. <https://www3.ha.org.hk/data/Provision/ApplicationProcedure>
- [26] Hkstp Ha Data Collaboration Lab[EB/OL]. [2025-06-08]. <https://www.hkstp.org/en/programmes/life-and-health-technology/hkstp-ha-data-collaboration-lab>
- [27] Hkstp Ha Data Collaboration Lab Workflow[EB/OL]. [2025-08-28]. <https://www.hkstp.org/-/media/corpsite/assets/programmes/institute-for-translation-research/hkstp-ha-data-collaboration-lab/hkstp-ha-data-collaboration-lab-workflow-v2.pdf>
- [28] 我国将建设健康医疗大数据产业体系[EB/OL]. (2016-06-18)[2025-02-23]. https://www.gov.cn/xinwen/2016-06/18/content_5083346.htm
- [29] 国家卫生健康委, 教育部, 科技部, 等. 关于印发涉及人的生命科学和医学研究伦理审查办法的通知(国卫科教发[2023]4号)[EB/OL]. (2023-02-18)[2025-08-28]. https://www.gov.cn/zhengce/zhengceku/2023-02/28/content_5743658.htm

[收稿日期:2025-09-01 修回日期:2025-10-11]

(编辑 赵晓娟)