

多中心生物样本资源共享与应用平台建设探索^{*}

郭 强

凌 鸿

杨 军 林伟龙

(中国医学科学院肿瘤医院 北京 100021) (神州医疗科技股份有限公司 北京 100000) (中国医学科学院肿瘤医院 北京 100021)

武志慧 闵浩巍 史文钊

时黎明 马建辉

(神州医疗科技股份有限公司 北京 100000) (中国医学科学院肿瘤医院 北京 100021)

〔摘要〕 介绍国内外生物样本库建设现状以及构建标准化生物样本库核心要素, 探讨多中心生物样本资源共享与应用平台建设, 包括平台基本框架、系统架构及功能模块, 分析生物样本库在临床研究中的作用。

〔关键词〕 生物样本库; 标准化; 信息化; 共享; 应用

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2021.03.012

Exploration on the Building of Multi-center Biological Sample Resource Sharing and Application Platform GUO Qiang, Cancer Hospital, Chinese Academy of Medical Sciences, Beijing 100021, China; LING Hong, Digital Health China Technologies Co., Ltd., Beijing 100000, China; YANG Jun, LIN Weilong, Cancer Hospital, Chinese Academy of Medical Sciences, Beijing 100021, China; WU Zhihui, MIN Haowei, SHI Wenzhao, Digital Health China Technologies Co., Ltd., Beijing 100000, China; SHI Liming, MA Jianhui, Cancer Hospital, Chinese Academy of Medical Sciences, Beijing 100021, China

〔Abstract〕 The paper introduces the building situation of biobank at home and abroad and the core elements of building standardized biobank, discusses the building of multi-center biological sample resource sharing and application platform, including the basic framework, system architecture and functional modules of the platform, and analyzes the role of biobank in clinical research.

〔Keywords〕 biobank; standardization; informatization; sharing; application

〔修回日期〕 2020-09-08

〔作者简介〕 郭强, 高级工程师, 发表论文 3 篇; 通讯作者: 马建辉, 主任医师。

〔基金项目〕 中国医学科学院医学与健康科技创新工程 (项目编号: 2017-I2M-2-003)。

1 引言

近年来临床诊疗数据和生物样本资源被广泛应用于病因研究、临床诊疗及疾病预防等领域, 优质生物样本资源是转化医学、精准医学快速发展的基础。我国人口多、地域广, 生物样本资源丰富, 如何制定适用性规范与标准, 实现生物样本资源整合

与共享、生物样本库可持续发展和科研价值最大化,是我国生物样本库建设过程中亟须解决的问题。本文通过建设多中心生物样本资源共享与应用平台,探索将现有生物样本库整合为具有统一建设标准、制度规范的生物样本库联盟的方法,以期为国家人口健康、疾病干预、科学研究、促进医疗服务可持续发展等相关政策制定提供科学依据。

2 国内外生物样本库建设现状

2.1 国外

目前国际上生物样本库建设模式主要分为两种:一是英国模式,即政府出资构建大规模、可满足不同需求的全国性生物样本库,典型代表包括英国生物样本库(UK Biobank)和丹麦国家生物样本库(Danish National Biobank);二是欧盟模式,即将现有生物样本资源进行整合,建立“分散实体库+统一信息库”的生物样本联盟网络,典型代表为泛欧洲生物样本库与生物分子资源研究中心-欧洲研究中心联盟(Biobanking and BioMolecular Resources Research Infrastructure - European Research Infrastructure Consortium, BBMRI-ERIC)^[1-3]。

2.2 国内

我国拥有庞大的临床生物样本资源,在政府大力支持下生物样本库发展迅速。自1994年建立中华民族永生细胞库以来相继建立了科室库、重大项目库、医院库和区域库等专项生物样本资源库,如脐带血造血干细胞库、上海张江生物银行、天津肿瘤医院肿瘤组织样本库、北京大学第一医院人类血液样本库、国家肝癌样本库及深圳国家基因库等^[4]。但由于建设理念、目标各异且相互之间协同性差,我国生物样本库建设情况与国际相比仍存在一定差距,主要表现为:样本资源分布不平衡,主要分布在具备科研资源优势的医疗机构,如北京、上海以及广州等城市的大型综合性医院内;样本类型重复集中,多数生物样本库基于某种科研需求而建立^[5]。此外还存在缺乏有效监管、尚未统一流程管理规范及可持续发展机制等问题^[6]。在保障安全的前提下

整合并共享现有生物样本资源,实现科研价值最大化是我国生物样本库建设工作的当务之急。

3 构建标准化生物样本库的核心要素与内容

3.1 核心要素

成功构建生物样本库离不开质量、信息和应用3个核心要素^[7]。生物样本库质量包括生物样本、信息、运作和管理质量^[7]。生物样本信息是指充分且特异性地应用临床、健康等其他相关信息描述(或注释)生物样本固有生物学特征^[8]。此外还应包括记录和管理样本采集、处理、出入库、存储和质量控制等信息。生物样本库应用是指将保存的生物样本用于疾病临床治疗和生命科学研究等,如探索药学领域和遗传治疗领域的新治疗手段,发现基因和环境对疾病的作用,制定个性化生活和环境方面的策略以进行疾病预防等。只有覆盖样本库运作、管理和服务的完整工作流程,才能实现生物样本资源的科学化管理和共享应用^[8]。

3.2 建设内容及意义

3.2.1 内容 包括实体库、样本信息库和样本资源信息共享平台。其中实体库负责样本保存,样本信息库用于存储样本的基本生物学特征、出入库等操作信息及相关临床诊疗信息(如患者、病史、检验、病理、影像检查、基因检测、治疗、随访等)。样本资源信息共享平台用来实现样本资源信息共享,支持科研产出。此外标准化生物样本库建设内容还包括搭建标准化样本库操作流程^[9]。国内外分别出台了相关标准和通用标准以规范生物样本库操作流程^[10-13]。

3.2.2 意义 有利于加强生物样本库建设与操作之间的协同性,减少不同生物样本库之间、生物样本与数据收集方式之间、相同主题信息与对应数据呈现方式之间的异质性,保障相同生物样本资源在不同时间进行操作与管理方式相同、不同生物样本库运作与管理方式相同,即保证不同生物样本库资源在相同主题下易于相互整合、交流与共享^[14]。

4 生物样本资源整合与共享

4.1 生物样本资源共享现状

4.1.1 定义及构成要素 生物样本资源共享是指按相关协议将生物样本库的生物样本和样本数据使用权交与其他生物样本库拥有，或与相关研究人员共同使用，为基础、临床和转化医学研究提供科学完整的生物样本、样本基本信息及关联信息。实现生物样本资源应用和共享是生物样本库存在的根本意义^[15]。生物样本资源共享包含 3 个构成要素：一是共享主体，即生物样本生产者（或提供者）和生物样本使用者（或需求者）；二是共享客体，即生物样本和信息资源，包括生物样本及其数量、种类、质量等信息；三是共享中介，包括共享制度、机制、法律、资金、平台技术等。共享主体、共享客体和共享中介之间以及各自内部关系的顺畅程度决定生物样本资源共享效果。

4.1.2 共享现状 我国生物样本库资源共享发展较晚，生物样本库整合和共享程度较低，主要存在样本类型重复但不全面、技术标准不统一、“私库”难以整合、样本质量参差不齐、样本信息未透明化、样本信息关联未完善、伦理规范匮乏、资源共享机制缺失等问题^[16-18]。

4.1.3 构建生物样本库资源共享机制 探索大数据时代背景下的生物样本库资源共享机制有助于实现人类生物样本的有序健康共享。我国目前尚无专门部门负责统筹和管理所有公共经费资助下的生物样本库建设和共享，参考其他国家生物样本库资源共享经验，可在某研究领域或区域构建生物样本库

联盟以实现生物样本库资源共享。需建立统一标准，规范共享内容和流程，搭建公共信息平台，确定各方认可的共享利益分配模式。上述机制可保证与生物样本库发生关联的各方都可以通过平台找到所需内容并保障多方收益，促进生物样本库持续运用和良好运转。

4.2 中国医学科学院肿瘤医院项目简介

4.2.1 基本情况 中国医学科学院肿瘤医院于 2019 年牵头“肿瘤生物样本库科研共享平台”建设工作。该项目以构建肿瘤生物样本资源科研共享平台为导向，开展基于样本资源内容与特征研究相关的生物样本库建设，通过以下 3 个方面实现生物样本的整合与共享：一是完善顶层设计，架构充分考虑系统易用性、扩展性和安全性。以浏览器/服务器（Browser/Server, B/S）结构为基础，利用自定义表单系统解决不同中心生物样本库之间数据异质性问题，实现对多中心样本信息的扩展及兼容，信息传输采用端到端加密方式保障数据安全。二是依照相关法律法规要求，设置生物样本库伦理委员会、执行委员会，针对知情同意、样本采集与使用申请、项目制定管理等方面进行伦理考核及审查。参照《生物样本库质量和能力通用要求》（GB/T 37864-2019），针对样本流通、样本可追溯性及信息与数据管理等方面进行标准化建立；三是探索开发科研项目管理模块，通过生物样本资源科研共享平台发布肿瘤领域研究课题，联合研究机构或人员开展临床相关研究，按入排标准筛选样本资源，真正实现多方应用和共享，见图 1。

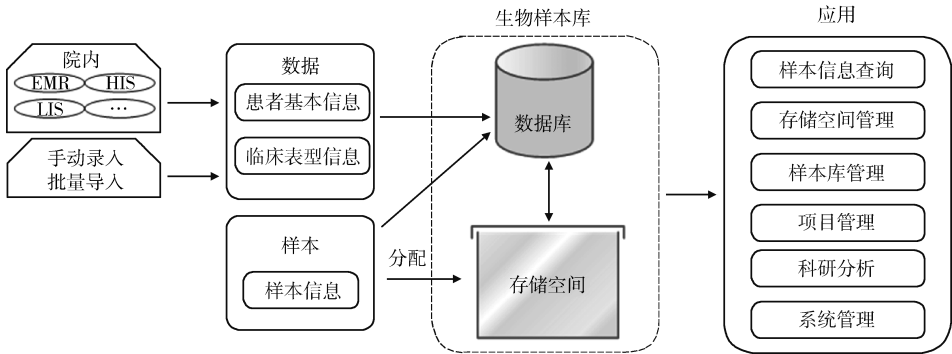


图 1 数据采集架构

4.2.2 基本框架 目前项目已初步完成院内5个分中心（肿瘤早筛、胸部肿瘤、结直肠癌、病理、肿瘤登记样本库）的资源整合与信息管理，能够对肿瘤样本的临床数据、样本基本信息及关联信息进行多维度管理，见图2。

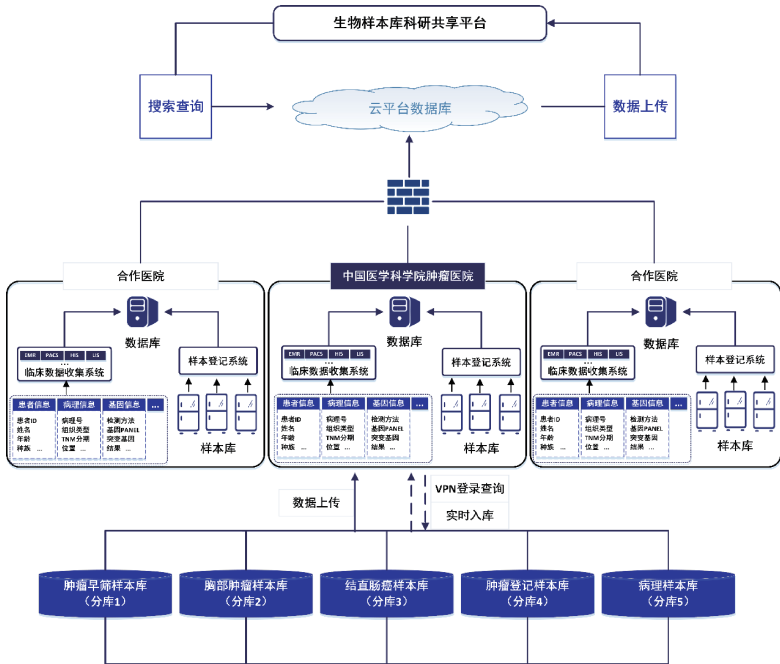


图2 肿瘤生物样本资源共享与应用平台基本框架

4.3 系统架构及功能模块

4.3.1 平台架构 构建多中心肿瘤生物样本库科研共享平台是一项体系化工程，涉及平台管理层、数据资源层、转化应用层、访问门户层、安全认证体系及标准体系。充分考虑系统易用性、扩展性和安全性，本平台架构为面向服务（Service - Oriented Architecture, SOA）的 B/S 架构。SOA 是一种面向复杂应用环境的框架设计与应用技术，将应用系统中所提供的功能和可访问的信息看作是一项服务，通过调用服务来获取功能，具有松耦合、位置透明和协议透明的优点^[19]。在此基础上，利用自定义表单系统解决不同中心生物样本库之间数据异质性问题，实现对多中心样本信息的扩展及兼容，信息传输采用端到端加密方式保障数据安全。

4.3.2 模块组成 （1）系统管理。妥善保管医疗数据是医院的首要职责，系统管理模块可实现域权限管理、用户管理和角色权限设置，通过对不同域下的不同用户分配角色来限定对应用和数据的访

问权限。（2）科研管理。实现用户对科研项目的管理，包括新建科研项目、定义科研项目入组标准、搜索符合入组标准的样本信息等，同时有助于其他用户以参与者身份参与科研项目，联合开展科研项目。（3）病种管理。实现生物样本库病种分类统计、病种图标展示等功能，用户根据需要创建、管理病种，设置病种名称、图标背景，与表单产生关联关系。（4）表单引擎管理。根据不同中心的需求自定义表单，实现各中心生物样本库个性化配置相关病种数据存储格式及类型，扩大应用场景，提高表单兼容性，解决不同中心不同样本库的数据异质性问题。（5）数据录入。信息录入功能中通过导航管理将表单与病种实现关联关系，根据用户权限信息录入，需要先选择病种再在系统录入相关信息。可选择病种通过当前用户权限下病种列表展示，录入信息为相关病种关联的表单元素，涵盖患者基本信息、就诊记录、现病史、家族史、基本检查、内镜检查/辅助检查、病理诊断、分子检测、药物治疗、疗效评估、临床症状、随访信息及样本库等。

上述信息可一次性录入完整也可分批录入。录入方式包括手动、自动导入。(6) 数据查询。实现基因数据信息、病理诊断结果等数据查询功能, 根据用户定义的统计分析模型可视化展示相关信息并生成统计分析报告, 以饼图、直方图和折线图等统计图表形式展示。查询到的感兴趣样本可联系相关样本持有人进行共享。

5 生物样本库在临床研究中的作用

5.1 为开展临床研究及实现精准医学奠定基础

生物样本库作为生物样本及数据资源存储载体, 包含样本相关的基本数据、病史信息、检验及检查信息、基因组学、蛋白组学及其他信息等, 是开展临床研究^[20]、实现精准医学^[21]的重要基础。

5.2 将患者基因与表型信息关联起来

为探索疾病发病机制和诊断方法提供便利。例如挖掘导致疾病发生的多个低风险致病基因突变信息需要借助生物样本库的大样本量组学研究^[22]。可通过代谢组学研究糖尿病、高血压等疾病, 这是因为对生物体液或组织中的代谢物进行检测、确定、定量和分类, 可表征继基因、mRNA 及蛋白等系列生物过程后的代谢产物, 以整体视角对生物体内代谢网络变化进行研究, 将代谢物与生物过程关联起来, 揭示机体在特定时间和环境下的整体状态^[23]。

5.3 用于生物标志物研究

生物标志物与人体生理、病理、变化密切相关, 从微观层面反映了病情变化和转归, 同时具有客观可量化特点, 在许多疾病的临床诊断或疗效评价中发挥重要作用^[24]。疾病的不同病程或不同样本来源的肿瘤标志物数值有所不同, 亟须寻找灵敏度、特异度较高的肿瘤生物标志物^[25]。

6 结语

我国拥有最大基数的生物样本资源, 生物样本为病因研究、新诊疗方法及新药临床研究等提供了

基础。标准化、智能化、规模化的生物样本资源管理平台能够实现生物样本资源的有效整合与共享。肿瘤生物样本库科研共享平台自建成以来已收集大量生物样本信息数据, 可对中国医学科学院肿瘤医院及合作单位开放共享权限, 实现样本资源共享。通过样本资源共享将宝贵资源应用于罕见病、肿瘤等重大疾病领域, 在基因、蛋白、免疫、细胞等方面开展科学创新, 加快科研成果转化为临床应用的速度, 同时将临床应用数据反馈给科研单位, 更加充分合理地利用样本资源, 推动基础研究, 为中国医疗事业做出贡献。

参考文献

- 1 史晓红, 郭健. 国际生物样本库的发展现状 [J]. 中华临床实验室管理电子杂志, 2017, 5 (1): 19-23.
- 2 方宇宁. 英国生物银行介绍 [J]. 世界科学, 2012 (9): 57.
- 3 Ommen G J B V, Outi Törnwall, Christian Bréchet, et al. BBMRI – ERIC as a Resource for Pharmaceutical and Life Science Industries: the development of biobank – based expert centres [J]. European Journal of Human Genetics, 2014, 23 (7): 893-900.
- 4 郜恒骏, 杜莉利, 张小燕, 等. 生物样本库发展的现状、机遇与挑战 [J]. 协和医学杂志, 2018, 9 (2): 172-176.
- 5 王庆宝. 生物样本库——转化医学与第六次科技革命 [J]. 泰山医学院学报, 2012, 33 (1): 1-3.
- 6 陈思静, 吴茂锋, 李佩娟. 生物样本库的建设与发展 [J]. 生物化工, 2019, 5 (4): 164-166.
- 7 满秋红, 王伟业. 重构生物样本库建设理念, 促进临床研究资源协同应用与发展 [J]. 协和医学杂志, 2019, 10 (4): 409-413.
- 8 王伟业, 周君梅, 蔡珍珍. 生物样本库的信息化管理与信息应用 [J]. 中华临床实验室管理电子杂志, 2017, 5 (1): 24-29.
- 9 杜莉利. 生物样本库的标准化建设 [J]. 转化医学杂志, 2016, 5 (6): 324-326.
- 10 《中国医药生物技术》编辑部. 中国医药生物技术协会生物样本库标准 (试行) [J]. 中国医药生物技术, 2011, 6 (1): 71-79.
- 11 中国标准化委员会. 国家标准委关于成立全国生物样本标准化技术委员会的批复 [EB/OL]. [2015-07-13]. <http://www.sac.gov.cn/szhywb/sytz/201507/t2015>

0713_190565.htm.

- 12 International Organization for Standardization. ISO20387: 2018 [EB/OL]. [2019 - 08 - 07]. <https://www.iso.org/standard/67888.html>.
- 13 国家市场监督管理总局 (国家标准化管理委员会). GB/T 37864 - 2019 生物样本库质量和能力通用要求 [EB/OL]. [2019 - 08 - 30]. <http://www.gb688.cn/bzgk/gb/newGbInfo?hcno=4FCF97DA42C358B2E83735834928F512>.
- 14 Quinlan PR, Mistry G, Bullbeck H, et al. A Data Standard for Sourcing Fit - for - purpose Biological Samples in an Integrated Virtual Network of Biobank [J]. Biopreserv Bioban, 2014, 12 (3): 184 - 191.
- 15 唐密. 生物样本库共享理论与实证研究 [D]. 上海: 复旦大学, 2014.
- 16 赵聪. 我国公共机构医学研究生物样本库共享问题研究 [D]. 北京: 北京协和医学院, 2017.
- 17 白莉华, 申锴, 杨军, 等. 生物样本库大数据的伦理与法律问题研究 [J]. 中国医学伦理学, 2017, 30 (10): 1206 - 1212.
- 18 雷瑞鹏, 冀朋, 冯君妍. 我国首部生物样本库伦理指南发布——“2017 年第二届中国生物样本库伦理论坛暨第

- 四届组学与生命伦理学研讨会”会议综述 [J]. 华中科技大学学报 (社会科学版), 2017, 31 (5): 139 - 140.
- 19 管红杰, 王珂, 江海峰, 等. SOA 架构的工作流管理系统的应用 [J]. 计算机工程与设计, 2011 (5): 1654 - 1657.
- 20 Sharma T, Guski LS, Freund N, et al. Suicidality and Aggression during Antidepressant Treatment: systematic review and meta - analyses based on clinical study reports [J]. BMJ, 2016 (352): i65.
- 21 Feero WG, Wicklund CA, Veenstra D. Precision Medicine, Genome Sequencing, and Improved Population Health [J]. JAMA, 2018, 319 (19): 1979 - 1980.
- 22 Hens K, Wright J, Dierickx K. Biobanks: oversight offers protection [J]. Science, 2009, 326 (5954): 798 - 799.
- 23 Nicholson J K, Lindon J C. Systems Biology: metabonomics [J]. Nature, 2008, 455 (7216): 1054 - 1056.
- 24 韩江, 周为民, 陈涵, 等. 如何发挥生物样本库在临床研究中的作用 [J]. 转化医学杂志, 2018, 7 (6): 346 - 349.
- 25 Myron G Best, Pieter Wesseling, Thomas Wurdinger. Tumor - educated Platelets as a Noninvasive Biomarker Source for Cancer Detection and Progression Monitoring [J]. Cancer Research, 2018, 78 (13): 3407 - 3412.

(上接第 58 页)

基于知识图谱的专家医生推荐方法, 以问答系统为基础构建患者需求模型, 通过数据挖掘构建基于诊疗专长的专家推荐模型, 将常规以搜索为核心转换为基于推荐方法的服务, 提高患者与专家匹配的准确度, 从而减少患者就医选择的从众性和盲目性, 降低就医的机会成本、时间成本和费用负担。精准的医患双向匹配模型不仅可以实现基于病症分析的精准就医, 同时也可以使专家医生匹配到可充分发挥自身专业优势的患者。

6.3.2 精准分时段预约 基于大数据技术和机器学习方法, 通过对历史诊疗数据的不断学习和反馈优化, 分析每名专家医生诊疗时间特点, 优化预约时段分配, 实现基于专家自身特点的精准分时段预约, 进一步缩短就诊等待时间, 动态调控专家号源数, 充分利用专家门诊时间。

参考文献

- 1 中华人民共和国国家卫生健康委员会办公厅. 关于加强

信息化支撑新型冠状病毒感染的肺炎疫情防控工作的通知 [EB/OL]. [2020 - 03 - 16]. http://www.gov.cn/zhengce/zhengceku/2020-02/05/content_5474692.htm.

- 2 马楠, 裴云庆, 魏国庆, 等. 大型综合医院以医疗为主导的“互联网+”实践探索 [J]. 中华医院管理杂志, 2017, 33 (2): 146 - 149.
- 3 浙江省卫生健康委员会. 关于印发 2019 年医疗卫生服务领域“最多跑一次”改革十大项目工作细则的通知 [EB/OL]. [2020 - 03 - 16]. http://www.zjwjw.gov.cn/art/2019/3/26/art_1202101_33410070.html.
- 4 World Health Organization. WHO Global Strategy on People - centred and Integrated Health Services: interim report [EB/OL]. [2020 - 03 - 16]. https://apps.who.int/iris/bitstream/handle/10665/155002/WHO_HIS_SDS_2015.6_eng.pdf?sequence=1&isAllowed=y.
- 5 钱辉, 张雪, 张大亮. 患者获取网络医疗信息风险的实证研究 [J]. 中华医院管理杂志, 2019, 35 (6): 494 - 498.
- 6 秦宇辰, 吴骋, 张新佑, 等. 国内外患者隐私保护立法情况对比及国外基本立法保护原则探析 [J]. 中国卫生事业管理, 2016 (1): 48 - 50.