转化医学研究（电子版）2011，1（2）：44－55
Translational Medicine Research（Electronic Edition），2011，1（2）：44－55

综　述



王晓民：教授，博士生导师。现任首都医科大学副校长，首都医科大学神经生物学系（国家重点学科）主任，北京市“脑重大疾病”重点实验室—省部共建国家重点实验室培育基地主任，教育部“神经变性病”重点实验室主任。目前主要从事帕金森病和阿尔茨海默病防治等研究。

按语：生物样本库是融合生物样本实体、生物分子信息以及样本表型数据的综合资源，对于开展人类疾病预测、诊断、治疗研究具有不可替代的重要作用。近年来国家支撑计划也立项支持在这一领域的创新研究。本文从生物样本类型、内容构成、样本规模、标准规范、质量控制、伦理保护、资源共享等方面与转化医学的关系在我国的发展情况和存在的问题，并对今后的发展提出了展望。相信读者通过这篇文章能更好地了解生物样本库与转化医学研究的重要性。

生物样本库与转化医学研究

张雷[1]，李海燕[2]，范可方[1]，韩秀丽[1]，胡纯严[1]，王晓民[2]*

1. 北京嘉和美康信息技术有限公司创新中心，北京 100085
2. 首都医科大学神经生物学系，首都医科大学北京神经科学研究所，神经变性病学教育部重点实验室，北京 100069

摘要　生物样本库是融合生物样本实体、生物分子信息以及样本表型数据的综合资源，对于开展人类疾病预测、诊断、治疗研究具有不可替代的重要作用。转化医学的兴起和发展对生物样本资源的迫切需求与日俱增，生物样本库从传统的单中心模式跃升到多中心网络化和国家级统筹发展的层次。作为支撑转化医学的重要基石，生物样本库需要进一步改进和完善，在生物样本类型、内容构成、样本规模、标准规范、质量控制、伦理保护、资源共享等方面满足转化医学的需求。

关键词　生物样本库，转化医学，转化研究

* 通讯作者：王晓民　E－mail：xmwang@ccmu.edu.cn

生物样本库（biobank）并不是一个非常新鲜的事物，其最早出现可以追溯到上百年前。从应用于疾病诊断、临床治疗，到应用于生命科学研究，人类社会不断发展的经济、文化、科技水平，也在不断赋予生物样本库更多的时代特色和日益丰富的内涵。

人类基因组计划（human genome project，HGP）的顺利完成，为人类健康与生命科学发展开启了一个新的阶段。通过对人类全基因序列、分类、序列多态性的解读，使从基因、蛋白等微观层面认知人类疾病的病因和转化机理，并探求个性化的、靶向性的诊断和治疗模式成为可能[1]。发现 biomarker，并将其有效运用到疾病的预防、诊断、治疗和药物研发过程，为患者提供个性化的疾病治疗方案，已经成为后基因时代生物医学研究的一个重要方向。快速发展的分子生物技术、生物信息技术，为基础医学研究提供了有力的技术支持，生命科学研究达到了前所未有的深度和广度，并不断催生越来越多的科学发现。但不幸的是，与这些资源耗费巨大的研究成果有关的临床验证工作却进行得非常缓慢[2]，大量的基础研究成果没有能够被快速、有效地转化到临床实践中来。为扭转这一被动局面，转化医学应运而生，并迅速发展成为生物医学研究领域备受关注的焦点。转化医学的关键在于实现从实验室（bench）到临床（bedside）的无缝隙集成与整合[3]，一方面能够将生物医学基础研究的最新成果快速有效地转化为创新的临床医学技术和产品（bench to bedside），另一方面将临床应用的实际情况快速反馈给实验室并开展针对性研究（bedside to bench），从而形成以临床需求为导向的正向研究循环。

生物样本库是集成基础研究与临床应用的重要技术载体，通过将临床表型数据与遗传数据进行有机整合，形成以研究个体为单位的生物医学信息资源平台，为转化医学发展提供重要的基础战略资源。在现代生物医学全面迈向分子医学、个性化医疗的这个时代，作为核心基础设施的生物样本库正迎来新的发展机遇和挑战。

1 定义

生物样本库，根据经济合作与发展组织（OECD）的定义，它是一种用于支持遗传研究的结构化资源，它既包括来自人体的生物标本实物（如血液、尿、组织、DNA、细胞等），也包括由对标本试验和分析过程产生的试验数据以及来与标本有关的大量关联信息（如人口统计学信息、临床诊疗信息、生活方式、环境信息等）[4]。不同的研究目的和研究设计都会对生物样本库的形态和内涵产生重要影响。但总体来说，生物样本库大致可以分为二种基本类型[5]：（1）人口生物样本库（population-based biobank）；主要面向随机选择的志愿人群，收集其生物标本以及家族史、生活方式、环境暴露因素等方面的数据资料。一方面可以用于研究常见疾病的自然发病情况和发展状况，另一方面还可以用于研究标志疾病发生的 biomarkers 和指示疾病倾向性的 biomarkers。（2）疾病生物样本库（disease-oriented biobank）；主要面向患病人群或健康对照人群，收集其生物标本以及各种临床表型数据资料。主要用于研究适于疾病诊断、疾病评估和疾病治疗的 biomarkers，发现新的药物治疗靶点。

转化医学（translational medicine）又称为转化研究（translational research），是国际医学领域近年来提出的一个新的学科分支。是指“从实验室到病床（bench to bedside）”和“从病床到实验室（bedside to bench）”的双向转化研究，简称为 B2B[6]。转化医学的实质是理论与实际相结合，基础与临床的整合，是多学科、多层次、微观与宏观、静态与动态、结构与功能、生理与病理、预防与治疗、人文与科学的交叉融合。

生物样本库是转化医学的重要基石、为发现和验证与疾病预测、预防、诊断、治疗有关的 biomarker 提供高质量的生物样本资源与技术服务。

2 生物样本库的国际发展状况

分子生物学、转化医学的快速发展，导致对大规模、高质量生物标本及相关信息资源需求的快速增长。因此，生物样本库的建设和应用得到世界很多发达国家政府、研究机构、生物医药企业的高度重视，将其视为发展生物医药领域核心竞争力的战略举措。

自上世纪 90 年代以来，各种类型的生物样本库如雨后春笋般出现在世界各地。据估计，在过去的十年里，美国用于生物样本库的建设经费已经达到 10 亿美元，在此期间，大约有 100 多个生物样本库建成，并收集了大约 345 000 例的生物样本资源[7]；同样，在欧洲、亚洲的一些主要发达国家，生物样本库也呈现快速发展的局面。2010 年，欧盟联合研究中心（JRC-IPTS）发布了一份针对欧洲现有生物样本库状况的调查报告，虽然其中的调查数据来自欧洲，

但在某种程度上也能够借以说明生物样本库在世界范围内的整体发展状况。在接受调查的来自 23 个国家的 126 个生物样本库中，77% 是在上世纪 90 年代以后开始兴建的，证明了生物样本库快速发展的态势；70% 是单中心模式；80% 由大学或政府负责管理，并向社会开放资源；85% 参与过多方合作研究；52% 涉及国际合作；大多数生物样本库的样本量属于中小规模水平，其中疾病生物样本库数量相对较多[8]。

生物样本库的广泛发展与应用虽然在一定程度上满足了某些特定研究的需求，但无法忽视的是，整个生物样本库领域的发展一直缺乏统一的标准支撑体系，涉及到术语、标识、数据结构、标准操作程序（SOP）、质量控制、资源共享、伦理、隐私与安全等多个方面。由于标准体系的缺失，给不同生物样本库间的资源交换与共享带来巨大的障碍。而现实中，很多大型的研究项目对样本规模的要求已经不是一个甚至几个生物样本库能够满足的，因此，解决生物样本库的资源孤岛问题已经成为其持续发展中无法回避的一个重大挑战。为应对这个挑战，一些国家开始建立国家级的生物样本库（表 1），以保障大规模的生物样本资源供应，这些行动代表着当前国际生物样本库发展的重要趋势。

表 1　国际生物样本库项目

名称	类型	启动时间	内容与规模	目　标
Icelandic Biobank 冰岛	人口样本库	1996	收集超过 14 万志愿者的生物标本、遗传数据和医疗数据。	在世界范围内，引导常见疾病遗传风险因素的研究，并将研究发现转化为临床诊断手段，用于评价个体患病风险。
Estonian Gene Bank 爱沙尼亚	人口样本库	2002	收集超过 100 万人的生物标本和健康信息（健康数据、家系数据、生活方式、环境因素、药物反应等）。	建立爱沙尼亚国家生物样本库，用于开展科学研究、遗传与健康研究，探索发现常见疾病的致病基因。
UK biobank 英国	人口样本库	2003	收集 50 万志愿者（40－69 岁的健康人群）的生物标本以及相关医疗信息。	研究生活方式、环境因素以及遗传因素对研究对象健康的影响，提高针对很多常见疾病的预防、诊断和治疗能力，促进社会整体健康水平。
Biobank Japan 日本	疾病样本库	2003	拟收集 30 万人的生物标本以及相关临床信息。	为研究个体基因差异，发展个性化医疗提供基础资源。
Hunt Biobank 挪威	人口样本库	2007	20 万人的全血和 DNA 样本；超过 10 万人的血清、血浆以及尿、RNA、细胞、血块黄层等样本；5 万人的肝素钠管。 样本关联数据包括：人口统计学、环境暴露、生活方式、健康体检、实验室检验、医疗数据（总计 3000 个变量，其中包含 800 个暴露变量）。	与国内以及国际的研究组织合作，应对世界当前面临的重要疾病挑战，如糖尿病、肿瘤、骨骼疾病、精神疾病、偏头痛、前列腺问题、尿失禁、生殖、心血管病等。 为 biomarker 验证，特定疾病的队列研究以及与卫生经济安全相关的市场后研究等提供数据支持。
caHUB 美国	疾病样本库	2009		加强研究基础，支持生物样本协作研究和基于循证的生物样本 SOP，实现生物样本库现代化。 开放结果数据、分析、策略文件以及科学工具，促进生物样本库收集能够满足特定科学需求的生物样本。
Lifegene 瑞典	人口样本库	2009	拟招募 50 万志愿者，收集血液和尿液样本以及志愿者的调查和健康检查数据。	深入理解基因、环境以及生活方式对健康的影响。

另外，一些国际化组织，包括：ISBER（International Society for Biological and Environmental Repositories）、P3G（Public Population Project in Genomics）、BBMRI（Biobanking and Biomolecular ResourcesResearch Infrastructure）、IARC（International Agency for Research on Cancer）、CaBIG（cancer Biomedical Informatics Grid）、OBBR（Office of Biorepositories and Biospecimen Research）等都在致力于推动生物样本库的统一协调发展，在生物样本库的最佳实践、技术标准、质量控制、资源共享、伦理规范等方面做了大量的研究工作，为生物样本库的持续健康发展奠定了坚实基础。可以预见，未来的生物样本库应该更加具有规模化、标准化、网络化和数字化的特点。

3 支持转化医学的生物样本库

大量的基础研究发现，要最终转化为临床可用的技术和产品，最少要通过两个阶段的验证工作。首先要确认这些成果是真实的，而不是由于统计方法导致的人工假象；然后还要通过前瞻性的临床研究，进一步确认基于这些基础研究成果的新型诊断或治疗方法能够切实提高患者的临床转归效果。在这两个验证阶段中，生物样本库都发挥着非常重要的作用，为整个转化过程提供不可或缺的生物样本资源和技术支撑工具。因此，生物样本库已经成为转化医学体系的重要基础平台（图 1）。

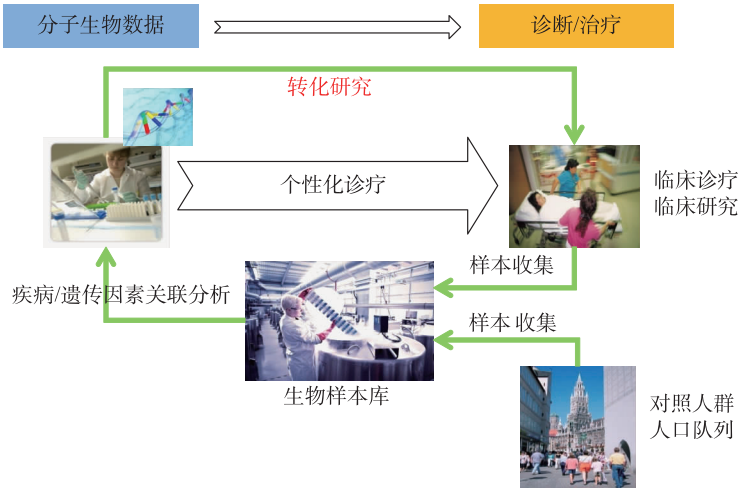

图 1　生物样本库与转化医学体系[9]

转化医学的快速发展，对生物样本库的建设和发展提出了更高的要求，这些要求涉及到生物样本库的类型、内容、规模、标准体系、质量控制、知情同意、共享机制等很多方面。

3.1 类型

转化医学不仅需要面向特定疾病的生物样本库，同时也需要面向人群的生物样本库。这两种类型的生物样本库为转化医学提供互补性的研究资源。基于面向特定疾病的生物样本库，研究者可以开展基因与疾病的关联性研究，以识别与特定疾病有关的基因表达模式；而基于面向人群的生物样本库，可以验证由全基因组关联研究产生的相关性发现。组合基于这两类生物样本库的研究，将有效加快转化验证进程。

3.2 内容

相比于传统的生物样本库，支持转化医学的生物样本库的内容更加丰富。它包括样本实体、表型信息、生物信息三个方面的内容，其中表型信息、生物信息的丰富程度直接决定了这个样本库的价值大小。表型信息一般应包括样本提供者全面的医疗信息、人口统计学信息、家庭信息、生活方式、职业信息、环境信息等等。由于不同研究设计的影响，表型信息的颗粒度往往会出现一定的伸缩性。以家族史为例，其具体内容可能会从单纯的父系、母系扩展到包含本代和子代的有关信息。生物信息是由实验室分析、试验衍生的生物分子数据、实验数据等，这些信息也应统一归属到样本库中，对这些数据的共享能够在一定程度上避免“重复发明轮子”的资源滥用问题。

3.3 规模

转化医学的兴起，对生物样本库的规模提出了极大的挑战。无论是基础研究发现还是其成果验证都需要使用大量的生物样本资源，较少的样本量会直接影响到各种试验分析的统计学结果，导致发现和验证在

真实性和可靠性产生偏差。传统的中小规模生物样本库已无法满足转化研究巨大的样本资源需求。样本资源的短缺，事实上已经成为制约转化研究的重要瓶颈。要实现生物样本库的规模化发展，有两种方案可供选择。一是建设超大规模的国家级生物样本库，这些样本库的规模可以达到几十万甚至上百万人的级别，如英国的 UK BIOBANK，瑞典的 LifeGene，爱沙尼亚的 Estonia Gene Bank 等；二是通过整合现有的样本库资源，形成集成、共享的网络化资源，如英国的 OnCore UK、美国的 caBig、加拿大的 CHRNet，欧洲的 P^3G 等。但是，一个超大规模的生物样本库从开始建设到可用状态，往往需要花费相当长的时间；而集成现有的样本资源又会面临着标准不一的协调问题。因此，采用并行方案可能是当前最现实的选择。

3.4 标准体系

转化研究要求所有的生物样本资源必须具有较好的一致性，以保证来自不同个体、不同机构、不同时间的样本及其相关数据之间的可比性，这是开展转化验证工作的重要前提。尤其随着一些大型医学研究项目的出现，不同的生物样本库最终会从孤立走向资源联合，以满足这些医学研究项目的样本需求。如果没有共同遵循的标准体系，上述资源联合将无法实现。目前，在生物样本库领域，还缺乏国际公认的标准体系，虽然一些国际组织（包括 WHO IARC、RAND、OECD、NCI 等）在致力于相关标准的研究工作，并发表了一些具有代表性的建设指南，但这些指南并没有也不可能对生物样本库进行全面的约束，每个样本库的建设过程都或多或少地融入了一些个性化的色彩，给不同生物样本库之间的资源共享带来潜在的风险。

生物样本库的标准体系建设是一项非常重要的基础工作，其中涉及到生物样本以及相关数据收集、处理、存储、应用的所有环节。

3.4.1 流程标准化

生物样本库的建设应遵循标准化的业务流程，通过流程的标准化，能够统一界定生物样本库建设的重要工作过程，明确每个过程的任务、执行者、控制条件以及接口约束，为整个标准体系的建立提供一个整体性的框架。

3.4.2 操作标准化

在生物样本库建设的每个业务过程，针对样本资源的流转与处理，应制定更加精细的操作标准（SOP）。如在知情同意环节，应制定统一的知情同意规范；在入组环节，应明确界定入组条件；在样本采集阶段，应全面定义样本采集的时间、部位、方法、工具等等；在信息收集过程，应确定内容范围、内容来源以及收集方式；在样本处理环节，应明确具体的操作资格、处理方法、设备、材料等等。只有共同遵循这些操作标准，才有可能保证生物样本各种状态的一致性。

3.4.3 数据标准化

数据标准化要求对生物样本相关的各种表型数据、生物分子数据的定义、表达、值域范围进行统一约定，以保证这些数据的一致性，在数据标准化方面应尽可能借引用国际上成熟的医学数据标准，如 SNOMED、ICD9/10、LIONC、DICOM 等。

3.5 质量控制

建立完善的质量控制体系是生物样本库建设不可缺少的重要基础工作。在很长一段时间内，这个问题并没有引起足够的重视[10]，各样本库所使用的 SOP 缺乏循证的支持，大家更多是在凭感觉，凭经验来进行标本的收集和处理，样本质量难以进行客观评价。

建立高水平的生物样本库质量控制体系的迫切要求加大生物样本质量保证方面的研究，正如美国 NCI OBBR（The NCI's Office of Biobanking and Biospecimen Research）以及欧洲 SPIDIA（Standardisation and improvement of generic Pre-analytical tools and procedures for In-vitro DIAgnostics）所做的那样，识别质量风险因素、完善质量追踪系统、建立循证的 SOP 以及量化质量评价手段（如组织分析、质量 biomarker 等）是保证生物样本质量的关键。

3.6 知情同意

与传统医学研究不同，生物样本库更多是面向未来的研究提供相关资源，对生物标本的处理和利用充满了不确定性[11]。因此，要实现传统意义上的知情同意，就变得不太现实。但过于开放的知情同意，又可能带来难以预料的伦理问题和争议。所以，分层的知情同意（由捐赠者事先选择样本的应用范围）或第三方监督下的开放式知情同意成为生物样本收集时可以考虑的两种选择。

3.7 共享机制

生物标本及数据资源的收集离不开临床工作者与研究者之间的紧密合作，大型转化研究项目也离不开不同生物样本库之间的资源合作，因此，有必要建立有关生物标本及其数据资源的共享机制。共享机制的设计应遵循“分层共享、有偿使用”的原则，通过量化的资源价值评估体系以及资源收益分配方式，让

临床工作者、研究者、资源提供者能够得到合理的回报，从而保证生物样本库持续滚动的发展，为转化研究提供源源不断的优质资源。

4 结语

在中国全面迈向创新型国家的征程中，中国生命科学与生物医药研究迎来新的发展机遇。在国内，转化医学虽然起步较晚，但发展迅猛。截止目前，已成立各级转化医学研究平台达到 30 多个。生物样本库作为整个转化医学体系的重要基石，其建设和发展应得到政府学界、企业的高度重视，通过借鉴、创新适合中国特色的生物样本库实践指南、质量控制体系以及资源共享机制，促进中国生物样本库达到更高的规模和质量水平，将中国丰富的临床病例资源真正转化成支撑健康与生命科学研究发展的无价之宝。

致谢：感谢北京重大疾病临床数据和样本资源库项目组给予的大力支持。

参考文献

[1] Expert Group on a National Cancer Biobank. Recommendations for the Establishment of a National Cancer Biobank. Dublin: © Health Research Board, 2008.
[2] President's Council of Advisors on Science and Technology. Priorities for Personalized Medicine. USA, 2008.
[3] Zimmerman Z, Swenson Ml, Reeve B. Biobanks: Accelerating molecular medicine challenges facing the global biobanking community. USA, © 2004 IDC, 2004.
[4] OECD. Guidelines on human biobanks and genetic research databases. France, © OECD, 2009.
[5] Martin A, Kurt Z. Biobanks: transnational, European and global networks. Briefings in Functional Genomics, 2007, 6 (3): 193 -201.
[6] Marincola FM. Translational medicine: a two-way road. J Transl Med, 2003, 1 (1): 1.
[7] Ákos Végvári, Welinder C, Lindberg H. et al. Biobank resources for future patient care: developments, principles and concepts. J Clin Bioinfor, 2011, 1 (1): 24.
[8] Zika E, Paci D, Schulte T. et al. Biobanks in Europe: prospects for harmonisation and networking. Spain, Publications Office of the European Union, 2010.
[9] Compton CC. The cancer human biobank (ca-HUB): advancing the vision of personalized medicine. 2nd Annual Biospecimen Research Symposium, USA, 2009.
[10] Malone B. NCI launches national biobank. Clin Lab News, 2010, 36 (3): 7.
[11] Hewitt RE. Biobanking: the foundation of personalized medicine, Curr Opin Oncol, 2011, 23: 112 -119.