

## **Biorepositories**



# ***Research White Paper***

---

## **Biorepositories**

### **Addendum to Registries for Evaluating Patient Outcomes: A User's Guide, Third Edition**

**Prepared for:**

Agency for Healthcare Research and Quality  
U.S. Department of Health and Human Services  
5600 Fishers Lane  
Rockville, MD 20857  
[www.ahrq.gov](http://www.ahrq.gov)

**Contract No. 290-2014-00004-C**

**Prepared by:**

L&M Policy Research, LLC  
Washington, DC

**Authors (alphabetical):**

Benjamin Greenberg, M.D., M.H.S. (Lead)  
University of Texas Southwestern Medical Center

Jennifer Christian, Pharm.D., M.P.H., Ph.D., FISPE  
IQVIA

Leslie Meltzer Henry, Ph.D., J.D., M.Sc.  
University of Maryland Law School; Johns Hopkins Berman Institute of Bioethics

Michelle Leavy, M.P.H.  
OM1

Helen Moore, Ph.D.  
National Cancer Institute, Biorepositories and Biospecimen Research Branch

**AHRQ Publication No. 17(18)-EHC014-EF**  
**February 2018**

The Effective Health Care Program of the Agency for Healthcare Research and Quality (AHRQ) conducts and supports research focused on the outcomes, effectiveness, comparative clinical effectiveness, and appropriateness of pharmaceuticals, devices, and health care services. More information on the Effective Health Care Program can be found at [www.effectivehealthcare.ahrq.gov](http://www.effectivehealthcare.ahrq.gov).

This report is based on research conducted by L&M Policy Research, LLC, with partners OM1 and IQVIA, under contract to the Agency for Healthcare Research and Quality, Rockville, MD (Contract No. 290-2014-00004-C). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

**None of the authors have any affiliations or financial involvement that conflicts with the material presented in this report.**

This report may be used and reprinted without permission except those copyrighted materials that are clearly noted in the report. Further reproduction of those copyrighted materials is prohibited without the express permission of copyright holders.

Persons using assistive technology may not be able to fully access information in this report. For assistance contact [EPC@ahrq.hhs.gov](mailto:EPC@ahrq.hhs.gov).

Suggested citation: Greenberg B, Christian J, Henry LM, Leavy MB, Moore H. Biorepositories. White Paper, addendum to Registries for Evaluating Patient Outcomes: A User's Guide, Third Edition. (Prepared by L&M Policy Research, LLC, under Contract No. 290-2014-00004-C.) AHRQ Publication No. 17(18)-EHC014-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2018. [www.effectivehealthcare.ahrq.gov](http://www.effectivehealthcare.ahrq.gov). DOI: <https://doi.org/10.23970/AHRQREGISTRIESBIO>.

**Reviewers (alphabetical):**

Nancy A. Dreyer, Ph.D., M.P.H., FISPE  
Chief of Scientific Affairs & Head Center  
for Advanced Evidence Generation  
IQVIA

Meg Ehm, Ph.D.  
GalaxoSmithKline

Richard E. Gliklich, M.D.  
Chief Executive Officer and Chairman  
OM1

Maria Harrison, B.S.  
PRA Health Sciences

Kathleen Mandziuk, M.P.H., R.N.  
PRA Health Sciences

## Contents

<b>Introduction.....</b>	<b>1</b>
<b>Patient Registries and Biorepositories .....</b>	<b>1</b>
<b>Types and Uses of Biorepositories and Their Application to Registries.....</b>	<b>2</b>
<b>Linking a Patient Registry and Biorepository.....</b>	<b>3</b>
<b>Development of a Biorepository .....</b>	<b>5</b>
Planning Considerations .....	5
<i>Scientific Purpose .....</i>	<i>5</i>
<i>Sample Collection and Storage.....</i>	<i>5</i>
<i>Oversight of Procedures .....</i>	<i>6</i>
<i>Business Model and Governance.....</i>	<i>6</i>
Operational Considerations.....	7
<i>Project Management.....</i>	<i>7</i>
<i>Quality Management.....</i>	<i>7</i>
<i>Sample Procurement.....</i>	<i>8</i>
<i>Sample Processing and Storage.....</i>	<i>8</i>
Linkage of Registry Data to a Biorepository .....	8
<b>Ethical and Regulatory Considerations.....</b>	<b>9</b>
Regulatory Considerations.....	9
Informed Consent .....	10
<b>Challenges in Developing and Using Biorepositories .....</b>	<b>17</b>
<b>Conclusions.....</b>	<b>18</b>
<b>References .....</b>	<b>19</b>

## Figures

Figure 1. Five approaches to informed consent and their correlation with the flexibility in future use of the collected data or biospecimens .....	12
--	----

## Introduction

The future of truly personalized medicine requires understanding each person's individual make-up and how that affects their risk of disease and response to treatments. Enormous scientific progress in understanding the genetic makeup of individuals has been made through analysis of biospecimens, and this knowledge is being used to conduct further research to better understand how diseases develop, progress, and respond to medical treatments. When properly collected, processed, and stored, biospecimens provide insight into the biology of the research participant, and in the case of individuals with disease, the biology of the disease. Thus, biologic specimens are an essential component to unlocking the interrelationships between biology and pathology.

Biorepositories can be either natural companions or integral to patient registries, and are being used to support a variety of research objectives. These data can enable longitudinal research studies to examine the biological and genetic risk factors for diseases along with the social, behavioral, and clinical factors that may be associated with the disease. For example, biospecimens collected over a research participant's lifespan can be meaningful tools to study aging, the development of disease, or progression or recurrence of disease when the genetic and biologic information is linked with other important determinants of disease.<sup>1-3</sup> Combining biospecimens with broader information about participant characteristics, in concert with longer term followup, may give insights into patterns that may activate or accelerate disease progress. Biospecimens collected at a single point in time can be used to identify biomarkers related to particular diseases or exposures or to support the development of new diagnostic tests. Further, it is important to keep in mind that some biospecimens are relevant descriptors of patients that are not anchored in time.

Although the linkage of patient registry data with biorepositories has great potential, there are many challenges that must be addressed in developing and operating a biorepository, as well as in linking data from biorepositories to registries. The purpose of this white paper is to provide an in-depth discussion of the use of biorepositories in medical care and clinical research, especially in the context of patient registries. This paper focuses on the use of biorepositories within the United States. Additional issues, including differences in regulatory environments, laws governing the import and export of samples, and cultural views on biospecimen collection, must be taken into account for repositories that collect biospecimens outside of the United States; a full discussion of these issues is beyond the scope of this paper. Where topics are well-covered in other materials, references and/or links are provided.

## Patient Registries and Biorepositories

Multiple terms may be used to refer to the collection and storage of biospecimens. For clarity, the most commonly used terms are defined below.

“Biospecimens” are human tissues, blood, and other bodily fluids removed from the body for medical or research purposes.<sup>4</sup> One of the most familiar examples of a biospecimen is blood removed from a person for a blood bank at a hospital or other medical organization, or donated by a person in the interest of helping others with a medical need for the blood. Other biospecimens might be removed from a person in the context of medical care, such as blood taken for routine metabolic tests, or blood, tissues, or other bodily fluids required for diagnosis of

a medical condition. When such samples are collected for medical care, there are sometimes extra biospecimens not needed for diagnosis. These biospecimens may be donated for research purposes by the patient. Biospecimens may also be donated for research in the context of research studies or clinical trials. Biospecimen donors are often referred to as research participants because of their critical role in supporting biological and medical research. In order to conduct such research, biospecimens from thousands and even millions of research participants are needed. It is important to note that the quality of biospecimen collection, processing, and storage, as well as the quality of the data associated with the biospecimen, are very important to the quality of the research conducted using those biospecimens.

“Biorepositories” are the entities that store biospecimens and distribute the samples to qualified researchers, an activity known as “biobanking.” The National Cancer Institute’s (NCI) Biorepositories and Biospecimen Research Branch developed a set of best practices to guide these entities, *The NCI Best Practices for Biospecimen Resources*,<sup>5</sup> first published in 2007 and updated in 2011 and 2016.<sup>6</sup> NCI uses the term “biospecimen resources” to cover all biobanking activities, from an individual researcher’s laboratory freezer to an industrial biorepository.

The scientific field of “biospecimen science” is the study of how different collection procedures affect the biological integrity of biospecimens and in turn, the quality and reproducibility of research utilizing biospecimens. NCI sponsors and conducts research studies in biospecimen science<sup>7</sup> and hosts a database of more than 2,000 publications in the field. The database, known as the “Biospecimen Research Database,”<sup>8</sup> also holds a library of hundreds of Standard Operating Procedures (SOPs) from around the world, all related to biospecimen handling. NCI and others are working to develop evidence-based biobanking practices to further improve the quality of biospecimen-based medical research.<sup>9</sup>

Lastly, a “patient registry” is defined as “an organized system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and that serves one or more predetermined scientific, clinical, or policy purposes.”<sup>10</sup> A patient registry may utilize biospecimens from a biorepository as a source of biological and genetic information and combine that information with other data sources, such as information from the patient, caregiver, or clinician, and diagnostic and/or treatment-related information.

### **Types and Uses of Biorepositories and Their Application to Registries**

Many types of biorepositories exist. Some are used to support research by biomarker validation, for example, whereas others contribute value by being linkable to registries or fully integrated with registries. The majority of biorepositories focus on collecting material from patients diagnosed with specific diseases. An example of a biomarker validation program is the Alzheimer’s Disease Neuroimaging Initiative, which uses biospecimens and clinical data collected from patients with Alzheimer’s disease or other forms of memory impairment to validate biomarkers for use in the diagnosis and study of Alzheimer’s disease clinical trials.<sup>11,12</sup> These same data, of course, are useful for registries. Other biorepositories are designed for the purpose of identifying genetic clues that can be used to guide therapeutic development.<sup>13-16</sup>

Closely akin to disease-focused biorepositories are those that focus on understanding specific habits and practices, such as playing sports. For example, a biorepository of brains from deceased professional football players in the United States is designed to learn more about the effects of this sport.<sup>17</sup>

A biorepository may be developed to support a single study, with specimens destroyed at the completion of that study. Alternately, a biorepository could be developed as a long-term resource for multiple research studies. An example of a biorepository that is supporting multiple studies is the Health Outreach Program for the Elderly (HOPE) at Boston University. The HOPE registry follows patients with Alzheimer's disease and normal controls on a yearly basis until death; after death, most participants donate their brain to the biorepository. Participants in the study are contacted regularly about participation in studies conducted at Boston University and elsewhere.<sup>18</sup>

Some biorepositories have an even broader purpose than a single disease. An example of a population-based biorepository with a broad focus is the UK Biobank, which was created to improve the prevention, diagnosis, and treatment of a wide range of serious and life-threatening illnesses, including cancer, heart diseases, stroke, diabetes, arthritis, osteoporosis, eye disorders, depression, and forms of dementia. Between 2006 and 2010, the UK Biobank recruited 500,000 people ages 40-69 years from across the United Kingdom. Participants have donated blood, urine, and saliva samples for future analysis, provided detailed information about themselves, and agreed to have their health followed over many years in an effort to help scientists discover why some people develop particular diseases and others do not.<sup>19</sup> Some efforts are taking the utility of biorepositories even further. For example, the eMERGE Network has focused on linking biorepositories to electronic medical records, which then can support many types of registries.<sup>20-22</sup>

Sponsors of biorepositories also vary. Some biorepositories are developed by medical centers or other organizations to collect and store biospecimens from a wide range of patients, while others are funded as part of specific research studies to collect specimens from study participants. Still others are organized by patient advocacy groups to spur research in a specific disease area.

In summary, biorepositories are used to support many different research purposes. Activities include medical research studies that have a specific disease focus, such as the identification of genetic mutations associated with a specific disease,<sup>23</sup> or more broadly-focused population studies, such as those utilizing the UK Biobank. Studies range from project-based studies of defined length and purpose, to ongoing studies in a particular disease area, to ongoing population studies.

### **Linking a Patient Registry and Biorepository**

While biorepositories are an essential tool in medical research, they often do not contain detailed clinical information or long-term followup information about the biospecimen donors. This limits the ability of the biorepository, in isolation, to examine, for example, the interplay of genetics and environment or changes resulting from therapeutic interventions. Patient registries typically contain detailed clinical information and, in some cases, long-term followup

information about participants. Combining this information with the genetic information in a biorepository can produce a valuable resource to support multiple research objectives.

The decision to link a patient registry with a biorepository should be guided, first and foremost, by the research objectives of the registry. The registry should have a clear purpose for linking the registry data to a biorepository. Beyond determining the research objectives, a critical first step is determining whether the registry will establish a new biorepository or link to an existing biorepository. Some registries establish new, dedicated biorepositories to support the research objectives of the registry. For example, the National Registry of Genetically Triggered Thoracic Aortic Aneurysms and Cardiovascular Conditions (GenTAC) was developed to support studies of the clinical management of genetically-triggered thoracic aortic aneurysms and related complications. The registry is a longitudinal, observational study that integrates genotypic, proteomic, clinical, and imaging data with outcomes data from approximately 3,000 patients enrolled at five centers across the United States.<sup>24</sup> GenTAC has supported multiple studies to date, including several focusing on genetic mutations associated with thoracic aortic aneurysms.<sup>25,26</sup> Another example is the Colon Cancer Family Registry (CCFR), an international consortium developed to facilitate studies of the genetic epidemiology of colon cancer. Participants at centers in the United States, Canada, Australia, and New Zealand provided blood samples and tumor tissue, along with demographic information, clinical history, and family history data. The registry recruited participants from 1997 to 2012, eventually enrolling 41,989 subjects who are recontacted every 5 years to complete followup questionnaires.<sup>27</sup> The registry has been used to support extensive genetic and molecular characterization work, resulting in over 270 publications to date.<sup>28</sup>

Setting up a biorepository specifically for the purposes of the registry has some advantages, namely the ability to control which biospecimens are obtained and the process for collecting, storing, and analyzing the specimens. However, establishment of a new biorepository can be costly and may be particularly complicated for registries that enroll participants in multiple countries. Alternatively, it may be possible in some scenarios to link registry data with an existing biorepository. One example is the linkage of the California Cancer Registry (CCR) with the University of California, Davis Cancer Center Biorepository (UCD CCB). Researchers performed a probabilistic data linkage between UCD CCB biospecimen records and CCR records and demonstrated that a large number of patients (81.2%) have information in both databases. Match rates for common cancers were high; for example, 93 percent of records matched for lung and respiratory system cancers, 91.7 percent matched for breast cancer, and 89.5 percent matched for colon and rectal cancers.<sup>29</sup>

This type of linkage between an existing biorepository and a patient registry – sometimes referred to as a ‘virtual biorepository’ – has the potential to support new research studies in an efficient manner. Other similar examples include the linkages between prospective cohort studies and cancer registries in the Nordic countries<sup>30</sup> and the linkage between the Surveillance, Epidemiology, and End Results Program (SEER) Residual Tissue Repository program and the population-based cancer registries in Hawaii, Iowa, and Los Angeles.<sup>31</sup>

Regardless of whether a registry builds its own biorepository or links to an existing biorepository, researchers who are considering linking a registry with a biorepository should understand the fundamental steps in planning and operating a biorepository. The following

sections describe the planning, operational, as well as regulatory, legal, and ethical issues that should be considered when developing a biorepository and/or using an existing biorepository to support new research projects.

## **Development of a Biorepository**

### **Planning Considerations**

#### ***Scientific Purpose***

The planning for a biorepository begins with an evaluation of the expected future use of the specimens and data collected. The variables that must be considered include whether the repository is created to support a hypothesis driven project or to enable a multitude of future research questions. Given the longitudinal nature of many patient registries, biorepositories linked to registries may be used to support several research objectives, as seen in the CCFR example described above. After the mission of the repository is established, developers must consider the population to be studied. Options include population-based cohorts (primarily used to study a variety of health conditions), health system-based cohorts (primarily used to study predefined pathologies) or condition-based cohorts (primarily used to create a resource for single disease-focused research). Additional types of cohorts may be specified by particular scientific questions to be answered within a biological or medical research project. Examples include the National Institutes of Health (NIH) Microbiome project<sup>32</sup> or the Genotype-Tissue Expression (GTEx) project,<sup>33</sup> both of which are studies that seek to understand “normal” human physiology and development. The dividing lines between these types of repositories can be blurred but must be considered in the initial planning stages.

#### ***Sample Collection and Storage***

The planning of the repository includes a determination of the sample types that will be acquired. To date, the majority of repository efforts have focused on genetic disorders and tumor banks, but there is a growing interest in repositories for a broad spectrum of conditions. Furthermore, the type of research being conducted on different diseases is crossing traditional disciplines. For example, the area of cancer immunology has forced traditional tumor banks to consider acquiring immunologically focused samples. The sample types have likewise grown in variety and complexity. Repositories currently collect tissue, whole blood, DNA, RNA, cells (e.g., PBMCs [peripheral blood mononuclear cells], fibroblasts), serum, plasma, cerebrospinal fluid, urine, nail clippings, hair, fecal material, or saliva.

Samples can be collected under dedicated research protocols or in the context of clinical care. In the setting of hypothesis-driven research, such as the collection of samples for a specific patient registry, the acquisition, processing and storage protocols are directed by the investigator utilizing the specimen in research. For other biorepositories, biospecimens are collected for future studies that have yet to be defined. Thus, repository developers need to design processes for acquisition, processing, and storage that conserve and protect the scientific quality of the specimen. SOPs for these specimens are part of the planning stage of repositories regardless of the mission or population being recruited and can be based on available protocols outlining best practices.

Based on the sample types to be collected and the number of patients, repository developers should plan for adequate storage needs. The storage facilities must include appropriate physical infrastructure, power, emergency power or cooling systems, monitored alarms, and the ability to track environmental conditions (e.g., temperature) for logging purposes. Critical to the planning stage of a repository is the need to plan for retrieval processes and potential expansion as recruitment grows.

### ***Oversight of Procedures***

Local oversight of a biorepository has several elements that must be considered during the planning stages. The oversight governs the collection, maintenance, and release of specimens *and/or* data. When designing a biorepository, a governance structure must begin with ownership of specimens. The entity that owns the samples and data may ultimately also be responsible for the conduct of the repository. Relative to collection of data and samples, the repository can either be the agent that interacts directly with the donors or the recipient of the data and samples from third parties. If the repository will be interacting directly with donors, then Health Insurance Portability and Accountability Act (HIPAA) regulations will have to be followed and an Institutional Review Board (IRB) is usually involved. The repository should outline procedures for informed consent, including who will complete the consent and where the consent form will be secured. The repository will also need to outline the procedures for acquisition of samples. For example, the sequence of tubes drawn during a blood collection needs to be determined and documented in advance as part of the SOPs, because some tubes cannot be drawn first in a sequence (e.g., Paxgene™ tubes). The development and use of evidence-based SOPs is encouraged so that SOPs are guided not by tradition or customary methods, but instead by the intended science to be accomplished. Oversight of these procedures is critical to ensure that sample quality is maintained throughout the life of the repository.

### ***Business Model and Governance***

Similarly, the repository requires a governance structure to monitor the storage and release of data and samples. If the repository is not created for a single investigator or patient registry but rather as a resource, then it will need a standard procedure for determining when and how to release data or specimens. Some repositories are operated on a fee schedule, while others are based on scientific review of the request. A clear governance structure must be defined at the initiation of the repository that specifies how and when samples and data will be released, as well as how research that uses the samples will be prioritized. The repository is responsible for maintaining a record of what samples and/or data is released, as well as when and to whom the samples are released.

As noted, the oversight procedures are based on the business model that is created for the repository. While the regulatory requirements for non-profit and for-profit repositories are the same, their governance models are distinctly different. For example, a non-profit repository may charge a fee for the use of data or specimens, but the fees typically do not cover the full costs of operating the biorepository.<sup>34</sup> Whichever model is used for a repository, the financial sustainability must be determined. Once a repository is initiated, there are fixed maintenance and upkeep costs that must be taken into account (e.g., storage, monitoring, quality controls, and oversight). Resources for better understanding the true costs of biobanking have recently become

available<sup>35</sup> and NCI has a freely available online tool for estimating biobank costs.<sup>36</sup> Furthermore, every repository should have an action plan for sample and data stewardship in case there is a significant loss of finances.

Project-based biospecimen collections, such as those associated with a specific patient registry, often have a firm end date for funding and plans must be made in advance for the disposition (transfer or destruction) of biospecimens. Recommendations for planning for these “legacy” stages of biobanking have recently been published.<sup>37</sup> Project-based biospecimen collections may be useful to researchers beyond the initial scientific goals, and if informed consent for broad future use has been obtained by participants the biospecimens can be made available through appropriate governance mechanisms including clear and transparent access policies. It can be challenging for biobanks to make their available biospecimens known to researchers. NIH has established some resources to enable such information sharing, including the NCI Specimen Resource Locator (SRL)<sup>38</sup> and the National Heart, Lung, and Blood Institute’s BioLINCC.<sup>39</sup>

### **Operational Considerations**

#### ***Project Management***

The overall management of a biorepository includes several roles that can be individualized or combined based on personnel numbers and expertise. A biorepository will need a programmatic manager, a financial manager, a supervisor to oversee data completeness and integrity, a technician to oversee sample processing, storage, and disbursement, and a governance board. In its most simple form, a biorepository could have a single researcher serve all of these roles, but as the size and complexity of a repository grows, then careful attention must be paid to each of these positions.

Project management is based on the underlying biorepository design. If the repository is a direct extension of a research study, such as a patient registry, then the repository management is part of the research study apparatus. If, however, the repository is being developed to support multiple, future research projects, then project management has a broader scope. In this case, the project manager of the repository is responsible for specimen collection and processing procedures, data collection processes and disbursement of specimens and/or data.

#### ***Quality Management***

Part of a repository program should be the inclusion of standard quality assurance procedures. These would include mechanisms for proper annotation to ensure sample integrity and data integrity. For biospecimens, quality control includes tracking the timeline of sample procurement, processing, and storage. Annotations should include the time and date stamp of sample procurement, processing, and when the sample was placed into storage. Some assays will be sensitive to the amount of time a biologic sample is *ex vivo* and as such, a repository needs to log this information. Secondly, the repository will need to set up standard quality control measures for different types of samples stored. For example, when storing PBMCs, it is standard to account for the cell viability post-processing and storage. Similarly, the integrity of the data collected is paramount. Having internal field validation processes should be standard in any

database. For example, a database should be able to validate the age of a person based on the date of data entry and the reported date of birth.

### ***Sample Procurement***

Sample procurement protocols will be based on the sample of interest and the clinical situation leading to sample acquisition. While there will be specific protocols for each type of specimen procurement (i.e., blood-based, tissue-based, and non-blood bodily fluid-based), these protocols will have to be incorporated into various scenarios for patient recruitment. Enrolling subjects into a repository tends to occur in one of three scenarios: a standalone process (i.e., research participant recruited to take part in a repository study); in the course of broader medical research; or in the course of routine clinical care. Each of these situations will lead to different procedures for procurement. For example, when obtaining specimens as part of a research project solely designed to build a repository, consideration only has to be given to the best protocol option for sample procurement that may serve a variety of future research uses. However, when incorporated into the context of routine clinical care, sample procurement procedures must account for other aspects of the research participant's situation. If obtaining tissue during a tumor resection, the process must consider the individual circumstances and operative variables when deciding how much tissue to resect. Only biospecimens that are not needed for patient diagnosis may be utilized for biobanking. When obtaining blood for a repository as an adjunct to a blood draw for clinical purposes, the total volume of blood obtained must account for the amount of specimen being obtained for clinical purposes as well as the research biospecimen. Each repository will need to consider these specific circumstances when developing procurement procedures.

### ***Sample Processing and Storage***

Once obtained, each specimen type should be processed based on best practice guidelines, but with consideration for future projects. For example, a repository may be built to collect specimens for a genetics study, and procedures may be developed to collect high quality DNA. Yet simultaneously, the serum from specimens could be isolated and stored for future projects on the same patient cohort. A broad, forward-looking approach to specimen processing and storage enormously increases the research value of the biospecimens and consequently, the value of the repository. Biorepositories when possible should choose biospecimen procedures that are optimal for the broadest range of potential future research use.

### ***Linkage of Registry Data to a Biorepository***

Biorepositories that will be linked to a patient registry require a mechanism to connect the prospective clinical data from an individual to all specimens collected from that individual. As noted above, linking this information dramatically increases the scientific value of the specimens obtained. A unique identifier may be used to link the specimens with the clinical data over time. It should be noted that many participants will have more than one sample and often more than one sample collected on the same date, which must be accounted for in the linkage mechanism. In addition, biorepositories that are developed to support a registry may collect additional samples over time that are linked to the same participant. Maintenance of patient privacy and confidentiality is an essential component of linking the samples and participants. As an example,

the Congenital Heart Disease Genetic Network Study maintains a biorepository linked to registry data. All clinical data are stripped of identifiers and labeled with a study number, while biospecimens are identified by a different number and stored securely in the biorepository. The study's administrative and data-coordinating center holds the link between the study numbers and the biospecimen identifiers.<sup>40</sup>

Inherent to this process of linking data is the need to obtain complete informed consent that allows for these procedures and explains the risks to each subject, as well as the need to comply with any applicable privacy regulations. These issues are discussed further in the following section.

## Ethical and Regulatory Considerations

### Regulatory Considerations

Biorepositories in the United States may need to comply with Department of Health and Human Services (DHHS), Food and Drug Administration (FDA), and HIPAA regulations, as well as ethical requirements. These regulations and their applicability to patient registry-based research are discussed at length in Chapter 7 of the User's Guide.<sup>10</sup>

Specifically relevant to biorepositories are the definitions of research. First, relative to privacy and HIPAA, DHHS regulations define *research*, at 45 CFR 46.102(d), as a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge. HHS regulations define *human subject*, at 45 CFR 46.102(f), as a living individual about whom an investigator conducting research obtains data through intervention or interaction with the individual, or identifiable private information. Private information must be individually identifiable (i.e., the identity of the subject is or may readily be ascertained by the investigator or associated with the information) in order for obtaining the information to be considered research involving human subjects.

In general, the Office for Human Research Protections (OHRP) considers private information or specimens to be individually identifiable, as defined at 45 CFR 46.102(f), when they can be linked to specific individuals by the investigator(s) either directly or indirectly through coding systems. Conversely, OHRP considers private information or specimens not to be individually identifiable when they cannot be linked to specific individuals by the investigator(s) either directly or indirectly through coding systems. For example, OHRP does not consider research involving **only** coded private information or specimens to involve human subjects, as defined under 45 CFR 46.102(f), if the private information or specimens were not collected specifically for the currently proposed research project through an interaction or intervention with living individuals; and the investigator(s) cannot readily ascertain the identity of the individual(s) to whom the coded private information or specimens pertain. This can be achieved by the investigators and the holder of the key entering into an agreement prohibiting the release of the key to the investigators under any circumstances, until the individuals are deceased.

The Privacy Rule is a Federal regulation under the HIPAA of 1996 (see 45 CFR part 160 and subparts A and E of part 164). The Privacy Rule permits covered entities under the Rule to determine that health information is de-identified even if the health information has been

assigned, and retains, a code or other means of record identification, provided that the code is not derived from or related to the information about the individual; the code could not be translated to identify the individual; and the covered entity under the Privacy Rule does not use or disclose the code for other purposes or disclose the mechanism for re-identification.

Maintenance of patient confidentiality is of particular concern for biorepositories, especially as more studies include DNA and RNA sequencing. Even though such data would be de-identified according to regulations, re-identification of research participants may be possible and has been reported in the scientific literature (see Chapter 16 of the User's Guide). Some research projects take additional steps to inform participants when the project will generate large volumes of data on individual research participants. For example, the NIH GTEx project requested informed consent from family members of the deceased when collecting biospecimens for GTEx, even though 45 CFR 46.102(f) does not define deceased individuals as human subjects. The individuals were de-identified, and the data derived from the project was posted in a database where only qualified researchers could obtain access.<sup>41</sup> As another example, the Congenital Heart Disease Genetic Network Study disclosed to all potential participants that the planned research included whole-genome analysis, and thus, it was theoretically possible that identifying information would be obtained.<sup>40</sup>

Lastly, relative to sample, handling, or processing, there are regulations that govern the safety of sample shipping and the requirements for handling samples with potentially infectious agents. The *NCI Best Practices for Biospecimen Resources*<sup>6</sup> and NCI Biospecimen Evidence-Based Practices<sup>9</sup> provide recommendations, not formal guidance or regulations, for federally funded research institutions. The College of American Pathologist's Biorepository Accreditation Program is voluntary at this time but can help biorepositories to achieve high quality operations.

Outside of the United States, regulations vary by country. A discussion of the regulatory considerations that apply to biobanking activities conducted outside of the United States is beyond the scope of this paper.

## **Informed Consent**

Obtaining voluntary informed consent from research participants has long been a cornerstone of ethically-conducted research.<sup>42</sup> The requirements for informed consent to federally-funded research are set forth in a set of 1991 regulations, often referred to as the "Common Rule," because 15 agencies, in addition to DHHS, have adopted them.<sup>43</sup> Until very recently, the Common Rule had not been amended since 2005. The evolution of patient registries and biorepositories posed challenges, however, to the classic understanding of informed consent. Large-scale databases and biorepositories create the opportunity to conduct future research with stored information and biospecimens. The range of possible future research, along with any risks or benefits associated with it, is generally unknowable when data and biospecimens are collected. At that juncture, researchers are typically unable to inform participants of all the ways in which their information and biospecimens may be used in the future, by whom, for what purposes, for how long, with what potential risks and benefits, and with what types of confidentiality protections. Against this background, concerns have arisen about whether researchers can comply with the conventional requirements for informed consent, and whether

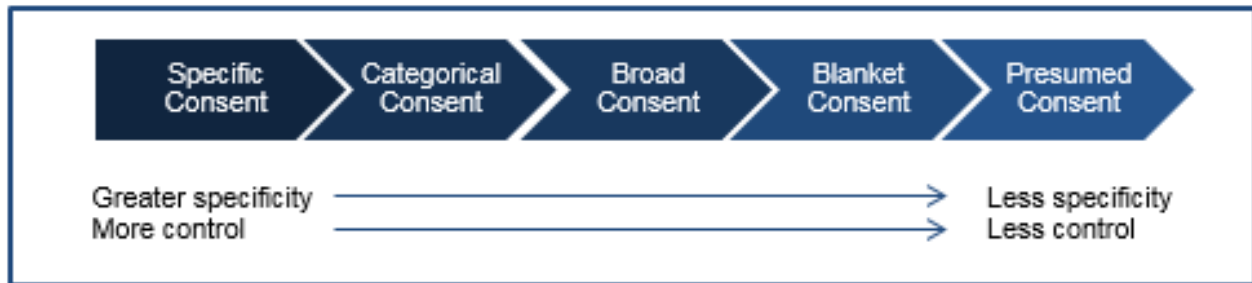
participants can make truly informed decisions about the future use of their data and biospecimens.<sup>44</sup>

In 2011, DHHS and the Office of Science and Technology Policy (OSTP) acknowledged that the regulations for informed consent had not kept pace with advances in biomedical research, including the proliferation of patient registries and biorepositories. Among the changes that DHHS and OSTP proposed in their advance notice of proposed rulemaking (ANPRM), entitled “Human Subjects Research Protections: Enhancing Protections for Research Subjects and Reducing Burden, Delay, and Ambiguity for Investigators,” was that written consent be required for the use of biospecimens in research, even if the biospecimens had been stripped of identifiers.<sup>45</sup> The 2015 Notice of Proposed Rulemaking (NPRM), which retained this provision, received numerous public comments, many of which were critical of that proposal. The Final Rule, enacted in 2017, is responsive to those concerns, and notably does not subject research involving non-identified biospecimens to the Common Rule.<sup>46</sup> The Final Rule does, however, contain provisions permitting investigators to obtain broad consent to the storage, maintenance, or secondary research use of identifiable private information and identifiable specimens. Specific elements of broad consent are discussed below.

As the Federal guidance has evolved, five models of informed consent for future use of data and biospecimens have been considered: specific consent, categorical consent, broad consent, blanket consent, and presumed consent. Specific consent involves obtaining consent from individuals for each discrete study in which their data and biospecimens will be used for research purposes. Categorical consent entails offering individuals a range of choices about how their information and biospecimens will be used in the future and by whom. Broad consent involves asking individuals to provide general consent to the future use of their data and biospecimens in research, subject to a few restrictions. Blanket consent means asking research participants to consent to *all* future uses of that material with no limitations on use. Presumed consent—which is sometimes referred to as an “opt-out” approach—assumes that all individuals who consent to the collection and storage of their data and biospecimens also consent to the future use of that material, unless they explicitly opt-out. These five approaches can be arranged along a spectrum in which there is a correlation between participants’ ability to control the future use of their data or biospecimens, and the specificity of the consent process (see Figure 1).

Although the 2017 Final Rule provides the option of broad consent for the storage, maintenance, or secondary research use of identifiable private information and identifiable specimens, it is important to understand all five approaches, particularly for research entities and investigators whose research falls outside Federal Common Rule oversight.

**Figure 1. Five approaches to informed consent and their correlation with the flexibility in future use of the collected data or biospecimens**



### *Specific Consent*

The specific-consent model would apply the classic doctrine of informed consent for research with human participants to the context of patient registries and biorepositories. The model starts from the premise that consent is a meaningful concept only when all of its traditional elements are satisfied with regard to a *specific* research study. According to this view, it is “impossible to obtain truly informed consent” to future research uses of data and biospecimens, because “the details that are a customary component of the traditional consent process cannot be disclosed.”<sup>47</sup>

The specific-consent model, therefore, would require researchers to recontact individuals and obtain their informed consent each time the researchers want to conduct new research on previously collected data and biospecimens that are linked to identifiable individuals. The model, which relies on the pre-2017 Common Rule governing informed consent to research, would require researchers to disclose the following informational elements to prospective participants when data or biospecimens are collected for storage and *each time* they are proposed for research use:

- Purpose and Procedures.** Researchers must inform prospective participants about the specific research study for which their data and biospecimens will be used, the scientific question to be answered, and the ultimate goal of the research. The consent process should also describe the procedures for collecting and storing the data and biospecimens. With regard to collection, participants should be told what procedures will be used to obtain biospecimens (e.g., a blood draw or biopsy); what kinds of information researchers intend to collect (e.g., demographic data, health history, or data in medical records); and how frequently the collection will occur (e.g., one-time only or at specific and ongoing intervals). With regard to storage, researchers should describe how data and biospecimens will be stored and labeled; where the data and biospecimens will be stored; whether it will be stored indefinitely, or respectfully destroyed when it is no longer useful for research; and whether it can be transferred to another storage resource in accordance with the terms of the informed consent.
- Procedures Related to Access, Sharing, and Recontact.** Researchers must also describe to participants who will be involved in the research, including who will have access to the data and biospecimens, whether large-scale data sharing will occur, and whether the researchers

intend to recontact the participants. Specifically, participants should be informed of the policies that will govern distribution of their data and biospecimens, including how access decisions will be made, what institutional body (e.g., ethics board) will review access requests; and what types of researchers will have access to the material, including commercial and foreign researchers.<sup>5</sup> If large-scale data sharing is envisioned, participants should be informed that their data and biospecimens may not remain in the original database or biorepository, and that different access policies may govern the distribution and sharing of those materials than the access policies specifically disclosed in the original consent document. Participants should also be informed that large-scale data-sharing can impact other aspects of the research study, including the potential risks of participation.

For researchers who might want to use the data and biospecimens they collect for future studies, the specific-consent model requires that they disclose, in the initial consent form, the procedures for recontacting participants. This information should include the variety of reasons researchers might have for recontacting participants after the primary study ends, how often recontact might occur, and the process for recontacting participants (e.g., mail, phone). Participants should be offered an opportunity to designate whether they consent to recontact or not. Under the specific consent model, if the participant does not consent to recontact, their identifiable data and biospecimens cannot be used in future research.

- ***Risks and Benefits.*** Researchers must inform prospective participants of any foreseeable risks and benefits of participating in the specific study proposed. The consent process should involve a description of the physical risks associated with collecting the biospecimens, as well as any potential benefits that could flow from the study to the individual participant or society at large. Researchers should also disclose that patient registries and biorepositories can pose risks to the participant's informational privacy. Emerging technology is making it increasingly possible to identify individuals even in de-identified, aggregated databases.<sup>48</sup> The unintended release or disclosure of personal information can lead to a variety of harms, including stigmatization, employment discrimination, and insurance loss.<sup>49-50</sup>
- ***Confidentiality.*** Best practices dictate that patient registries and biorepositories establish clear procedures for protecting the confidentiality of participants' identifiable information.<sup>5</sup> These procedures may include coding, encryption, limited access to data and biosamples, and non-disclosure agreements. They may also include NIH-issued Certificates of Confidentiality, which can protect particularly sensitive information from forced disclosure.<sup>51</sup> Researchers must describe these protections to participants as part of the informed consent process, but they should not guarantee confidentiality. Rather, they should inform participants that, despite substantial efforts to safeguard their confidentiality, a chance remains that their confidentiality may be breached.
- ***Financial Considerations.*** Researchers must disclose any costs that the participant may incur as a consequence of their participation in the research study. If biospecimens will be collected in the course of routine clinical care, for example, participants should be informed of whether they are responsible for the costs associated with the collection procedure. Researchers should also disclose the possibility that their research may result in the creation

of new commercial products. Participants should be informed of whether they will share in any research-related profits.

- ***Voluntary Participation and Withdrawal.*** Researchers must inform participants that their participation in the study is voluntary, that refusal to participate will not result in a penalty or loss of any benefits to which they are otherwise entitled, and that they may discontinue participation at any time without penalty or loss of any benefits to which they are otherwise entitled. In the context of patient registries and biorepositories, it is important at the outset to establish a policy to address participant withdrawal. To date, registries and biorepositories have adopted various approaches, ranging from completely removing the participants' data and biospecimens to prohibiting future use of that material, but allowing current research to continue. Regardless of what policy is followed, any limitations on complete withdrawal should be disclosed to participants during the consent process.

Although specific consent is considered the gold standard for ensuring that research participation is fully informed and voluntary, some commentators have argued that in the context of large patient registries and biorepositories, it is impractical. One concern is that obtaining specific consent for each study is administratively difficult, time consuming, and costly.<sup>52</sup> Another concern is that participants may be lost during followup or may decline future participation when they are recontacted, either of which can jeopardize the quantity and quality of future research.<sup>53</sup> These factors have led a number of commentators to argue in favor of alternatives to specific consent for future research use of data and biospecimens.

### *Categorical Consent*

Categorical consent is an alternative to specific consent that offers participants an opportunity to choose, from a menu of options, how their data and biospecimens will be used in the future. This approach—sometimes referred to as “tiered consent” or “checklist consent”—takes place during the initial consent process to data and/or biospecimens collection and research. In addition to providing consent to the initial research study, participants would be offered the option of consenting to certain types of future research, which may be categorized by type of research (e.g., disease or condition)<sup>54</sup> or type of researcher (e.g., academic, government, commercial, or international).<sup>55</sup> Under this model, participants' data and biospecimens may not be used for future research that falls outside of the categories initially approved by the participant. For example, the GenTAC project provided participants with a tiered set of options regarding the use of their biospecimens. Eleven percent of participants refused permission to create a cell line, 12 percent refused to allow commercial access to their DNA, 7 percent refused to allow outside researchers access to any type of sample, and 6 percent refused to have their samples stored indefinitely.<sup>56</sup>

The primary benefit of categorical consent is that allows participants to exercise some autonomy with regard to how and by whom their data and biospecimens will be used in the future. Participants can agree to future research that aligns with their values and preferences, and they can withhold consent from future research to which they object (e.g., mental health research, cloning, for-profit research). Critics of categorical consent argue, however, that it is not “informed consent” because participants do not know the specific purposes or risks associated with future studies to which they agree to participate. As one commentator has noted, “the more

general the consent is, the less informed it becomes.”<sup>57</sup> A second drawback of categorical consent is that to operate effectively, it requires systems that can reliably detect and honor participants’ choices. Maintaining those systems can be financially and administratively burdensome for registries and biorepositories.<sup>58</sup>

### *Broad Consent*

Broad consent, as discussed above, refers to a process by which individuals prospectively consent to the future use of their data and biospecimens for a broad and unspecified range of biomedical research. Researchers may exclude certain types of future research from the broad consent if there is evidence that they may be objectionable to a large number of people (e.g., human cloning) or if there are scientific reasons to restrict use of certain biospecimens (e.g., limiting the use of biospecimens to study people with a rare disease), but generally, there are few exceptions to broad consent.<sup>59</sup>

The 2017 Final Rule, which adopts the broad consent model, stipulates that broad consent must satisfy the general informed consent requirements in the Common Rule, as well as the following additional elements for any research involving the storage, maintenance, or secondary research use of identifiable private information and identifiable specimens<sup>46</sup>:

- A general description of the types of research that may be conducted with identifiable private information or identifiable biospecimens;
- A description of the identifiable private information or identifiable biospecimens that might be used in research, whether sharing of such information or biospecimens might occur, and the types of institutions or investigators that might conduct research with such information and biospecimens;
- A description of the period of time allowed that the identifiable private information or identifiable biospecimens may be stored or maintained, and a description of the period of time such information or biospecimens may be used for research purposes;
- Unless the subject or legally authorized representative will be provided details about specific research studies, a statement that they will not be informed of the details of any specific research studies that might be conducted using the subject's identifiable private information or identifiable biospecimens, including the purposes of the research and that might have chosen not to consent to some of those specific research studies;
- Unless it is known that clinically relevant research results, including individual research results, will be disclosed to the subject in all circumstances, a statement that such results may not be disclosed to the subject; and
- An explanation of whom to contact for answers to questions about the subject's rights about storage and use of the subject's identifiable private information or identifiable biospecimens, and whom to contact in the event of a research-related harm.

Importantly, the Final Rule does not require institutions to adopt broad consent, or what is sometimes referred to as “one-time general consent.”<sup>60</sup> The revised Common Rule offers investigators the *option* of obtaining broad consent for future, unspecified research use of identifiable private information and identifiable specimens. Additionally, under the revised rule, if an individual is asked to provide broad consent and declines, then the IRB may not permit waiver of consent.<sup>46</sup>

The Final Rule aligns with empirical data demonstrating that while most people want to control whether their data and biospecimens are used in future research, the majority of people are willing to broadly consent to research on any medical condition, without recontact for each specific study.<sup>53, 61, 62</sup> Nevertheless, critics of broad consent maintain that “any consent to future research projects that are not clearly described, is by definition invalid because it is not informed.”<sup>44</sup>

### *Blanket Consent*

Blanket consent involves asking individuals to prospectively consent to *all* future research with their data and biospecimens, without any restrictions. The primary argument advanced for blanket consent is that it respects participants’ autonomy without impeding socially beneficial research. Participants are informed about the wide range of ways their data and biospecimens may be used in future research, and if they are uncomfortable with the open-ended research parameters, or feel they do not align with their personal values, they can choose not to participate. Some surveys suggest that this “yes or no” approach to consent is consistent with public preferences: in one study, when given a range of choices about how their biospecimens would be used, most research participants either consented to unlimited future research or no research at all.<sup>63</sup>

Critics of blanket consent argue, however, that it privileges scientific progress and the public good over long-standing ethical commitments to respect human research participants and their autonomy.<sup>47</sup> Participants are asked to consent to future research with almost no information about, and no control over, how and by whom their data and biospecimens will be used. One commentator has argued that blanket consent poses “non-welfare” harms to participants when their materials are used in future studies that conflict with their fundamental values.<sup>63</sup> Other commentators suggest that blanket consent may lead to the overconsumption of biospecimens for commercial, technical, or forensic applications, to the detriment of biomedical research.<sup>53</sup>

### *Presumed Consent*

Under a presumed consent model, researchers assume that all participants who consent to the collection and storage of their data and biospecimens also consent to the future research use of that material unless they proactively opt out. This approach, which is sometimes referred to as “opt-out consent,” shifts the informed consent presumption from one in which individuals are included in research only if they consent, to one in which they are included in research unless they refuse.

Proponents of presumed consent contend that it not only overcomes the high costs and efforts associated with obtaining informed consent for each research use, it also increases research

participation by changing the default from exclusion to inclusion.<sup>52</sup> Critics of presumed consent argue, however, that it does not satisfy the requirements for informed consent, because participants are not told, and do not consent to, the purpose, risks, and benefits of each research use of their data and biospecimens. Commentators have also raised concerns about the administrative difficulty of adequately informing all potential participants the opportunity to opt out and the enrollment in research of people who do not wish to participate but fail to opt out.<sup>64</sup>

### **Challenges in Developing and Using Biorepositories**

While the development of biorepositories and the linkage of biorepositories and patient registries has great potential, these efforts also face several challenges. First, appropriate and legal informed consent is critical for biobanking activities, but additional mechanisms to engage research participants in biobanking are under discussion. These include active public education on biobanking so that the public better understands the critical role of individual donations in furthering medical research; an emphasis on intentionally asking patients to become research participants, rather than employing opt-out approaches for biobanking in medical institutions; and considering the return of research results to participants, when appropriate and upon agreement with participants and researchers. NCI has developed patient brochures to better communicate with potential research participants<sup>65</sup> and video communications to better explain the context for biospecimen donation.<sup>66</sup>

Harmonization of biospecimen collection, processing, and storage procedures according to evidence-based guidelines is also needed. Biospecimens are collected in medical laboratories all over the United States and the world, often using different procedures for many of the critical steps in biospecimen collection, and very often not recording information about these steps in a manner where that data can accompany the biospecimen. These critical steps can include: obtaining proper informed consent of the research participant and keeping appropriate records of that consent; expedient handling and processing of the biospecimens; and proper storage and distribution conditions. An example of where harmonization is important is when researchers are looking for biological indicators, or biomarkers of a disease process. The researchers may need to look at biospecimens from hundreds or thousands of research participants to find the biomarkers. If the researchers are using biospecimens that were collected in very different ways, the biological processes may be altered in the biospecimens, making it more difficult to find the “common denominator” biomarkers that hold the key to the disease process. NCI and other groups are actively working in this area to both develop the scientific data needed for evidence-based practices and to develop and disseminate the practices. The *NCI Best Practices for Biospecimen Resources* was developed to begin a process of improving research biospecimen quality and, in turn, improving the quality and reproducibility of research results. The document provides baseline recommendations for operations and encourages harmonization of biospecimen practices across medical institutions to facilitate appropriate research sharing of biospecimens and increase the quality of research results. In addition, the College of American Pathologists has recently launched a program to accredit biobanks using a peer-based accreditation system, based on the *NCI Best Practices for Biospecimen Resources* and other available information.

In addition to harmonization, collaboration between medical institutions and researchers to share difficult-to-find biospecimens is critical to accelerate research. Medical institutions and

researchers in the past have often considered the biospecimens in their possession to be for their exclusive use, but as more biospecimens are needed for research, there has been an increased emphasis on the need for sharing and collaboration. In addition, biospecimens collected specifically for a clinical trial, patient registry, or other study could be deposited in biorepositories for future use (assuming appropriate patient consent is obtained), rather than being destroyed at the conclusion of the study.<sup>67</sup>

Lastly, improved understanding of the economics of biobanking is needed. As stated above, many different models for biobanking are in use today, from limited, project-based biobanking to broad biobanking programs that span years and even decades. It is important to ensure that sustainable funding is available for biobanking and that biobanks focus efforts on financial planning and appropriate cost recovery mechanisms to stay solvent. NCI has recently conducted two surveys on biobanking economics<sup>35</sup> and has developed an online tool for biobank financial planning.<sup>36</sup>

### **Conclusions**

In summary, biobanking activities are an essential component of biological and medical research. Linkage of biorepositories with the detailed clinical and often longitudinal data in patient registries can produce a valuable resource for medical research, but multiple challenges must be addressed for the linkage to be successful. As best practices continue to evolve and as requirements for informed consent are clarified, it is likely that interest in linking biorepositories and patient registries will grow, particularly as the concept of personalized medicine increases in importance.

## References

1. Genetic Alliance Registry & BioBank. <http://www.biobank.org/>. Accessed January 10, 2017.
2. Susan G. Komen Tissue Bank at the IU Simon Cancer Center. <http://komentissuebank.iu.edu/>. Accessed January 10, 2017.
3. Army of Women: A Program of the Dr. Susan Love Research Foundation. <http://www.armyofwomen.org/drlovefoundation>. Accessed January 10, 2017.
4. Biorepositories and Biospecimen Research Branch, Patient Corner: What are Biospecimens and Biorepositories. <http://biospecimens.cancer.gov/patientcorner/default.asp>. Accessed January 10, 2017.
5. Office of Biorepositories and Biospecimen Research, National Cancer Institute. NCI Best Practices for Biospecimen Research. 2011.
6. Biorepositories and Biospecimen Research Branch: Best Practices. <http://biospecimens.cancer.gov/bestpractices/>. Accessed January 10, 2017.
7. Biorepositories and Biospecimen Research Branch: Biospecimen Research Network. <http://biospecimens.cancer.gov/researchnetwork/>. Accessed January 10, 2017.
8. Biorepositories and Biospecimen Research Branch: Biospecimen Research Database. <https://brd.nci.nih.gov/brd/>. Accessed January 10, 2017.
9. Biorepositories and Biospecimen Research Branch: NCI Biospecimen Evidence-Based Practices. <http://biospecimens.cancer.gov/resources/bebp.asp>. Accessed January 10, 2017.
10. Gliklich R, Dreyer N, Leavy M, eds. Registries for Evaluating Patient Outcomes: A User's Guide. Third edition. Two volumes. (Prepared by the Outcome DEcide Center [Outcome Sciences, Inc., a Quintiles company] under Contract No. 290 2005 00351 TO7.) AHRQ Publication No. 13(14)-EHC111. Rockville, MD: Agency for Healthcare Research and Quality. April 2014. <http://www.effectivehealthcare.ahrq.gov/registries-guide-3.cfm>. Accessed January 10, 2017.
11. Alzheimer's Disease Neuroimaging Initiative. <http://www.adni-info.org/Scientists/ADNIOverview.html>. Accessed January 10, 2017.
12. Barnes J, Bartlett JW, Fox NC, et al. Targeted recruitment using cerebrospinal fluid biomarkers: implications for Alzheimer's disease therapeutic trials. *J Alzheimers Dis.* 2013;34(2):431-7.
13. Naj AC, Jun G, Beecham GW, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet.* 2011;43(5):436-41.
14. Reitz C, Jun G, Naj AC, et al. Variants in the ATP-binding cassette transporter (ABCA7), apolipoprotein E ε4, and the risk of late-onset Alzheimer disease in African Americans. *JAMA.* 2013;309(14):1483-92.
15. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet.* 2013;45(12):1452-8.
16. Wang LS, Naj AC, Graham RR, et al. Rarity of the Alzheimer disease-protective APP A673T variant in the United States. *JAMA Neurol.* 2015;72(2):209-16.
17. Boston University CTE Center. <http://www.bu.edu/cte/our-research/brain-donation-registry/>. Accessed January 10, 2017.
18. Boston University Alzheimer's Disease Center. Clinical Core Overview. <http://www.bu.edu/alzresearch/about-us/our-cores/clinical/>. Accessed January 10, 2017.

19. UK Biobank. <http://www.ukbiobank.ac.uk/>. Accessed January 10, 2017.
20. McCarty CA, Chisholm RL, Chute CG, et al. The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BMC Med Genomics*. 2011;4:13.
21. Gottesman O, Kuivaniemi H, Tromp G, et al. The Electronic Medical Records and Genomics (eMERGE) Network: past, present, and future. *Genet Med*. 2013;15(10):761-71.
22. Smoller JW, Karlson EW, Green RC, et al. An eMERGE Clinical Center at Partners Personalized Medicine. *J Pers Med*. 2016;6(1):pii: E5.
23. Yurgelun MB, Masciari S, Joshi VA, et al. Germline TP53 Mutations in Patients With Early-Onset Colorectal Cancer in the Colon Cancer Family Registry. *JAMA Oncol*. 2015;1(2):214-21.
24. Eagle KA, GenTAC Consortium. Rationale and design of the National Registry of Genetically Triggered Thoracic Aortic Aneurysms and Cardiovascular Conditions (GenTAC). *Am Heart J*. 2009;157(2):319-26.
25. Roman MJ, Devereux RB, Niles NW, et al. Aortic root dilatation as a cause of isolated, severe aortic regurgitation. Prevalence, clinical and echocardiographic patterns, and relation to left ventricular hypertrophy and function. *Ann Intern Med*. 1987;106(6):800-7.
26. Silverman DI, Burton KJ, Gray J, et al. Life expectancy in the Marfan syndrome. *Am J Cardiol*. 1995;75(2):157-60.
27. Newcomb PA, Baron J, Cotterchio M, et al. Colon Cancer Family Registry: an international resource for studies of the genetic epidemiology of colon cancer. *Cancer Epidemiol Biomarkers Prev*. 2007;16(11):2331-43.
28. National Cancer Institute. Colon Cancer Family Registry Cohort (CCFRC). <http://epi.grants.cancer.gov/Consortia/members/ccfr.html>. Accessed March 23, 2016.
29. McCusker ME, Cress RD, Allen M, et al. Feasibility of Linking Population-Based Cancer Registries and Cancer Center Biorepositories. *Biopreserv Biobank*. 2012;10(5):416-20.
30. Pukkala E, Andersen A, Berglund G, et al. Nordic biological specimen banks as basis for studies of cancer causes and control--more than 2 million sample donors, 25 million person years and 100,000 prospective cancers. *Acta Oncol*. 2007;46(3):286-307.
31. SEER Residual Tissue Repository Program. [Internet]. Bethesda (MD): National Cancer Institute; 2011. Sep 21, 2011. <http://seer.cancer.gov/biospecimen>. Accessed January 10, 2017.
32. Data Analysis and Coordination Center: NIH Human Microbiome Project. <http://hmpdacc.org/>. Accessed January 10, 2017.
33. NIH National Human Genome Research Institute, The Genotype-Tissue Expression Project (GTEx). <http://www.genome.gov/27543767>. Accessed January 10, 2017.
34. Vaught J, Kelly A, Hewitt R. A review of international biobanks and networks: success factors and key benchmarks. *Biopreserv Biobank*. 2009;7(3):143-50.
35. Odeh H, Miranda L, Rao A, et al. The Biobank Economic Modeling Tool (BEMT): Online financial planning to facilitate biobank sustainability. *Biopreservation and Biobanking*. 2015;13(6).
36. Biorepositories and Biospecimen Research Branch: Biobank Economics Modeling Tool. <http://biospecimens.cancer.gov/resources/bemt.asp>. Accessed January 10, 2017.
37. Matzke L, Fombonne B, Watson P, et al. Fundamental Considerations for Biobank Legacy Planning. *Biopreservation and Biobanking*. 2016;14(2):99-106.

38. National Cancer Institute: Specimen Resource Locator. <https://specimens.cancer.gov/tissue/default.htm>. Accessed January 10, 2017.
39. National Heart, Lung, and Blood Institute, BioLINCC | Biologic Specimen and Data Repository Information Coordinating Center. <https://biolincc.nhlbi.nih.gov/home/>. Accessed January 10, 2017.
40. Pediatric Cardiac Genomics Consortium. The Congenital Heart Disease Genetic Network Study: rationale, design, and early results. *Circ Res*. 2013 Feb 15;112(4):698-706.
41. Carithers LJ, Ardlie K, Barcus M, et al. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreserv Biobank*. 2015 Oct;13(5):311-9.
42. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects Research. Washington, D. C.: United States Department of Health and Human Services 1979.
43. 45 CFR 46.101(a). <https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html#46.101>. Accessed January 10, 2017.
44. Elger B, Biller-Andorno N, Mauron A, et al. Consent and Use of Samples. *Ethical Issues in Governing Biobanks: Global Perspectives*. Burlington, VT: Ashgate Pub. Ltd.; 2008. pp. 57-88.
45. Human Subjects Research Protections: Enhancing Protections for Research Subjects and Reducing Burden, Delay, and Ambiguity for Investigators, 76 Fed. Reg. 143 (proposed July 26, 2011).
46. Federal Policy for the Protection of Human Subjects; Final Rule. 82 Federal Register 12; 19 January 2017; 7149-7274.
47. Caulfield T. Biobanks and Blanket Consent: The Proper Place of the Public Good and Public Perception Rationales. *Kings Law Journal*. 2007;18:209-26.
48. Gymrek M, McGuire AL, Golan D, et al. Identifying Personal Genomes by Surname Inference. *Science*. 2013;339(6117):321-4.
49. Eriksson S, Helgesson G. Potential harms, anonymization, and the right to withdraw consent to biobank research. *Eur J Hum Genet*. 2005 Sep;13(9):1071-6. PMID: 15986039.
50. Fullerton SM, Lee SS. Secondary Uses and the Governance of De-Identified Data: Lessons from the Human Genome Diversity Panel. *BMC Med Ethics*. 2011;12:16.
51. National Institutes of Health. NIH Announces Statement on Certificates of Confidentiality. NOT-OD-02-037. <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-02-037.html>2002. Accessed January 10, 2017.
52. Rosenbloom ST. Ethical and Practical Challenges to Studying Patients Who Opt Out of Large-Scale Biorepository Research. *J Am Med Inform Assoc*. 2013;20:e221-e5.
53. Hansson MG. Should Donors Be Allowed to Give Broad Consent to Future Biobank Research? *Lancet Oncology*. 2006;7:266-9.
54. National Bioethics Advisory Commission. *Research Involving Human Biological Materials: Ethical Issues and Policy Guidance*. Rockville, MD: 1999.
55. Beskow LM, Dean E. Informed consent for biorepositories: assessing prospective participants' understanding and opinions. *Cancer Epidemiol Biomarkers Prev*. 2008 Jun;17(6):1440-51. PMID: 18559560.
56. Kroner BL, Tolunay HE, Basson CT, et al. The National Registry of Genetically Triggered Thoracic Aortic Aneurysms and Cardiovascular Conditions (GenTAC): results from phase I and scientific opportunities in phase II. *Am Heart J*. 2011 Oct;162(4):627-32.e1.

57. Arnason V. Coding and Consent: Moral Challenges of the Database Project in Iceland. *Bioethics*. 2004;18(1):27-49.
58. Vaught J, Lockhart NC. The evolution of biobanking best practices. *Clin Chim Acta*. 2012 Oct 9;413(19-20):1569-75. PMID: 22579478. PMCID: 3409343.
59. Grady C, L. E, Berkman B, et al. Broad Consent for Research with Biological Samples: Workshop Conclusions. *American Journal of Bioethics*. 2015;15(9):34-42.
60. Wendler D. Consent for Research with Biological Samples: One-Time General Consent Versus a Gift Model. *Ann Intern Med*. 2012;156(8):598-.
61. Brothers KB, Morrison DR, Clayton EW. Two large-scale surveys on community attitudes toward an opt-out biobank. *Am J Med Genet A*. 2011 Dec;155A(12):2982-90. PMID: 22065592. PMCID: 3222722.
62. Chen DT, Rosenstein DL, Muthappan P, et al. Research with stored biological samples: what do research participants want? *Arch Intern Med*. 2005 Mar 28;165(6):652-5. PMID: 15795341.
63. Tomlinson T. Respecting donors to biobank research. *Hastings Cent Rep*. 2013 Jan-Feb;43(1):41-7. PMID: 23254911.
64. Kaufman D, Bollinger J, Dvoskin R, et al. Preferences for opt-in and opt-out enrollment and consent models in biobank research: a national survey of Veterans Administration patients. *Genet Med*. 2012 Sep;14(9):787-94. PMID: 22538255.
65. Biorepositories and Biospecimen Research Branch, Public Resources | Brochures and Educational Materials | How You Can Help Medical Research: Donating Your Blood, Tissue, and Other Samples. <http://biospecimens.cancer.gov/resources/publications/educational.asp>. Accessed January 10, 2017.
66. Biorepositories and Biospecimen Research Branch, Multimedia | Donating Your Tissue for Research <http://cdp.cancer.gov/resources/reports/multimedia.htm>. Accessed January 10, 2017.
67. Mora M, Angelini C, Bignami F, et al. The EuroBioBank Network: 10 years of hands-on experience of collaborative, transnational biobanking for rare diseases. *Eur J Hum Genet*. 2015 Sep;23(9):1116-23.