Establishing a Cancer Center Data Bank and Biorepository for Multidisciplinary Research

Christine B. Ambrosone,1,2 Mary K. Nesline,1,2 and Warren Davis2
1Department of Epidemiology and 2Division of Cancer Prevention and Population Sciences, Roswell Park Cancer Institute, Buffalo, New York

Abstract
The establishment of a biorepository with linkage to clinical and epidemiologic data will provide an invaluable resource for cancer research, including studies of cancer etiology, progression, and prognosis, as well as development of biomarkers for early detection. Developing an infrastructure for a biorepository linked to clinical, pathologic, and epidemiologic data requires significant efforts in strategic planning for efficient means to ascertain, identify, and consent participants, as well as guidelines for blood collection, processing, and storage while maintaining participant privacy rights. In this report, we present an approach to developing a Data Bank and Biorepository at our own institution, with discussion of elements to be considered when establishing such a bank. (Cancer Epidemiol Biomarkers Prev 2006;15(9):1575–7)

Introduction
Explosive progress has been made in identifying molecular alterations that occur in the multiple steps and pathways in the carcinogenic process, and the elucidation of the human genome has led to promising research on the effect of genetic variants on cancer risk and prognosis. The usefulness of basic science advances, however, may be limited without the ability to evaluate molecular characteristics in relation to patient information, age, race, gender, disease characteristics, exposure data, and importantly, recurrence and survival status, particularly in light of therapies received. Once available, linkage of molecular data to demographic, pathologic, and clinical records, enhanced by epidemiologic data, will greatly facilitate studies of cancer risk, progression, and outcome in an environment that maintains strict confidentiality. Those who contribute such data must provide informed consent, and their rights to participate, decline to participate, and withdraw must be protected. In this article, we describe our approach to the establishment of a Data Bank and Biorepository (DBBR), which unites all elements as a resource for the conduct of research based on human subjects.

Establishing the Infrastructure for a DBBR
The establishment of a DBBR requires decision making on several fronts. Key questions to be addressed include:

(a) Who will be eligible for participation?
(b) How will potential participants be ascertained, approached, and consented?
(c) How will blood samples be obtained?
(d) How will epidemiologic information be collected?
(e) How will participant information be tracked, and what data will be collected?
(f) What types of tubes will be used for collection, and how will blood be processed and stored and
(g) How will clinical data be linked with specimens and epidemiologic data?

The answers to these questions are dependent on the foreseeable uses of data and samples from the DBBR and are likely to vary depending on the needs of the potential users and the structure of the institution.

Eligibility. Our DBBR is clinic based, with standardized recruitment and enrollment procedures applied within each outpatient clinic (breast, gynecologic, colorectal, thoracic, genitourinary, etc.). To some extent, patient eligibility in each of the clinics is driven by the needs of the members of the Disease-Specific Working Groups, composed of clinicians, pathologists, and research scientists. The overall flow for the DBBR is shown in Fig. 1.

Patients with Cancer. To ensure that only patients who will be followed through our institution would be enrolled and not those seeking second opinions, our cancer patient population includes only those who have been diagnosed with cancer at our institution and are scheduled for treatment and follow-up.

High-Risk Patients. Patients who are at high risk of cancer and are being seen by the Clinical Genetics service are also eligible for participation, with a focus on women and men at high risk of breast and colon cancer. Additionally, women who are seen in breast clinic and meet the Gail model criteria for high-risk breast cancer are asked to participate in the DBBR.

Healthy Controls. In case/control studies where cases are ascertained at a hospital or Cancer Center, controls should derive from the same population as the cases and be likely to be treated at the same institution if they were diagnosed with cancer. Thus, we enroll visitors or family members of patients who are approached for participation in the DBBR. These noncancer participants can be matched by sex, age, county of
residence, and race to serve as “controls” for patients with other types of cancers. For example, women accompanying men being seen for prostate cancer can serve as controls for women with breast cancer.

**Recruitment.** Potential participants are identified by daily review of clinic schedules, and those likely to be eligible are checked against records of patients already enrolled or those who have refused. Communication between DBBR research associates and clinic staff ensures that patients will be seen for the informed consent procedure at a convenient time and before going to phlebotomy. This procedure allows for the collection of blood specimens before the receipt of any treatment, including surgery, and minimizes additional venipuncture for research purposes.

The consent includes separate check-off boxes for permission for the collection of biological specimens (blood, buccal cells, sputum, and urine), the completion of an epidemiologic questionnaire, follow-up, and recontact for future studies, and linkage of all data and specimens with medical records for demographic, clinical, and laboratory test results and with results from tumor tissue studies. Permission to use surgical tissue for research is part of the surgical consent process.

When consent is completed for the patient and any accompanying family members or friends, the research associate provides participants with an enrollment packet that includes the bar-coded DBBR questionnaire and self-addressed postage-paid envelope for return, a copy of the consent form, and an Institutional Review Board–approved DBBR brochure.

Following informed consent, the research associate logs into the patient database of the institution and enters orders for the bloods to be drawn for the DBBR, which are signed by authorized clinic personnel. The procedures for this have been established with the Clinical Laboratory Medicine Department. As shown in Fig. 2, when the participant proceeds to phlebotomy, blood is drawn into the appropriate tubes, labeled with a hospital system-generated barcode ID number, and sent through the pneumatic tube system to the station located in proximity to the DBBR laboratory. Bloods are centrifuged into plasma, RBC, buffy coat, and serum and are aliquoted into 0.5 mL straws using the CryoBio System MAPI (Paris, France). Aliquots of whole blood are cryopreserved for storage of viable cells for future immortalization. Laboratory personnel scan the barcode into the liquid nitrogen freezer software, which tells the staff the location (in the tank) of where specimens will be stored. The time from blood draw to storage in liquid nitrogen is limited to 1 hour or less, to maintain integrity of blood samples and to minimize variability in markers that could be due to sample degradation.

The DBBR as a data warehouse consists of three databases:

(a) A secure, permission-restricted consent-tracking database contains participants’ contact and basic demographic information, consents, and database linking identifiers (unique IDs).
(b) The blood processing laboratory separately maintains the specimen storage and retrieval database. This database is

---

**Figure 1.** Schema for infrastructure for DBBR.
populated during specimen processing using barcode IDs only. Unlike the consent-tracking database, it does not contain any participant protected health information and all specimens are aliquoted and stored by barcode ID.

(c) The DBBR separately maintains the epidemiologic questionnaire database. This database is populated by scanning participants’ returned bar-coded questionnaires. Like the specimen-tracking database, the questionnaire database does not contain any protected health information. DBBR also links to a separate database for abstracted medical record data, tumor registry data, data acquired from disease-specific clinical registries, and follow-up event data.

These databases are not maintained as one “merged” database to enable Health Insurance Portability and Accountability Act compliance. The DBBR data manager queries the databases using unique IDs to fulfill approved requests for biological material and associated epidemiologic questionnaire data. The records supplied to investigators contain barcode ID only and will never include the key contained in the consent-tracking database.

Applications for use of specimens and data are reviewed by the Steering Committee members of the site-specific Disease-Specific Working Groups. Banked specimens are available for researchers whose protocols have been approved by the institutional Institutional Review Board and prioritized by the Disease-Specific Working Groups Committee. Specimens and data are distributed to investigators with no identifying participant information attached. It is also requested that laboratory results from samples obtained from the DBBR be returned for entry into the DBBR database for use by other investigators so that assays will not be done in duplicate by other researchers. These criteria will facilitate collaborative, multidisciplinary relationships among investigators.

Future Directions

This article provides a brief overview of the necessary procedures for establishment of a DBBR to serve as a resource for cancer research. The utility of a biospecimen bank, with specimens collected from patients who have consented to linkage of samples with clinical and epidemiologic data as well as results from tissue studies, cannot be underestimated. However, the procedures necessary to collect data and samples from patients who have provided informed consent for their use is not a trivial matter. It is important for those wishing to establish a DBBR to consider the questions posed at the beginning of this article and to let the potential uses of the resource guide its establishment. Clearly, determination of the participants to be included in the bank and the types of specimens that will be stored will be influenced by the needs of the potential users. However, there are some concepts that have universal importance, including the necessity for procedures to protect patient confidentiality and maintain compliance with Health Insurance Portability and Accountability Act, standardization of time from blood draw to processing and storage, and establishment of detailed tracking systems (1). As the numbers of samples increase and assays done on those samples generate thousands of data points to be managed, the need for tracking databases will require extensive efforts in bioinformatics.

Reference

Establishing a Cancer Center Data Bank and Biorepository for Multidisciplinary Research

Christine B. Ambrosone, Mary K. Nesline and Warren Davis


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/15/9/1575

Supplementary Material
Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2006/09/22/15.9.1575.DC1

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.