The Biobank at The University of Texas Health Science Center at Houston

Jennifer E. Sanner and Krystle J. Nomie

Introduction

THE UNIVERSITY OF TEXAS HEALTH SCIENCE CENTER at Houston (UTHealth) Center for Clinical and Translational Sciences (CCTS) Biobank was initially created using the biospecimens and related clinical data collected from the Texas Medical Center Genetics (TexGen) Project. The TexGen project was started to promote and aid medical and genetic research conducted in the Texas Medical Center primarily focused on cardiovascular diseases as well as cancer. In forming the UTHealth CCTS Biobank, the disease categories were expanded to also include more disease categories and minority representation. The ultimate goal of the UTHealth CCTS Biobank is to promote translational research via the sharing of biospecimens and related clinical data to interested researchers. In addition, the UTHealth CCTS Biobank has begun implementing policies and processes to share genomic data.

Infrastructure: The biobank uses the federated model of sharing in which the contributor maintains ownership of the donated biospecimens and related clinical data.¹ Furthermore, under this model, the biospecimens and related clinical data remain with and are governed by the contributing principal investigators (PI). The contributing PI/ or owner has agreed to share the biospecimens and related clinical data based upon scientific merit, inventory availability, and correlation with their studies. All biospecimens and related clinical data within the UTHealth CCTS Biobank have been previously consented to secondary use by the study participants.

Biospecimen types: plasma, red blood cells, peripheral blood mononuclear cells (PBMCs), serum, and tissue

Standard operating procedures (SOPs): SOPs have been established and implemented for the sharing of biospecimens and related clinical data and data management processes. Currently, SOPs are being developed regarding the acquisition and sharing of genotypic data.

Sample location enhanced distribution (SLED) web application: The UTHealth Biobank in collaboration with the UTHealth CCTS Biomedical Informatics group designed and implemented an online search web application termed SLED (Sample Location and Enhanced Distribution). This i2b2-based application allows interested researchers to search the UTHealth inventory for biospecimens and related clinical

data that are applicable for their studies. Contributors to the Biobank have provided a minimal data set pertaining to their biospecimens with no associated protected health information (PHI) to ensure that the patient identity remains secure. These de-identified datasets are used as the data source for the CCTS Biobank i2b2 Extract-Translate-Load (ETL) process and are then uploaded into the online SLED interface to enable the ability to search these data sets. In addition, SLED is able to capture and upload the requestor's application documents to maintain the files in an easily accessible account associated with the requestor. Overall, SLED acts in place of the manual process previously implemented to search for biospecimens and related clinical data as well as manages requestor applications.

Start date of operations: September 2006

Category: clinical/translational/basic

Focus of biobank: The primary disease categories included in the UTHealth CCTS Biobank are cardiovascular disease (including acute coronary syndrome and premature acute myocardial infarction); cancer, inflammatory diseases (primarily sclerosis), stroke, aortic aneurysms, and diabetes mellitus (primarily type-II diabetes mellitus). In addition, the UTHealth CCTS Biobank strives to be representative of Texas, which has a large Hispanic population, representing the changing landscape of the United States where minority groups are increasing. We are currently recruiting contributors primarily working with Hispanic populations to join the Biobank. Hispanics currently represent 20% of the UTHealth CCTS Biobank participant population.

Funding sources: Clinical and Translational Science Award (CTSA; UL1 TR000371) from the National Institutes of Health with a voluntary institutional match of 100%. The CTSA was first granted in September 2006 and was renewed in June 2012.

Metrics:

- A total of approximately 14,000 biospecimens distributed since 2006
- As of 2014, approximately 160,000 biospecimens currently available
- As of 2014, the genomes of approximately 7,000 participants have undergone whole exome sequencing (WES)

- Researchers/projects assisted: 47 with 5 in the pipeline as of January 2014
- Publications: 51 since 2006

Most valuable/highest priority cases: Since the Biobank would like to be more representative of Texas and the changing national landscape, our Hispanic samples are very high priority. In addition, the Biobank would like increase the sharing of the WES data (7000 patients); therefore, we are strongly marketing these data to interested researchers.

Major challenges: Recruiting interested researchers and contributors, implementing a cost recovery plan

Major milestones: Launch of SLED

Reference

1. Vaught J, Kelly A, Hewitt, R. A review of international biobanks and networks: Success factors and key benchmarks. *Biopreserv Biobank* 2009;7:143–150.

Address correspondence to:
Jennifer E. Sanner
6901 Bertner Avenue, Suite 567
Houston, Texas 77030

E-mail: Jennifer.E.Sanner@uth.tmc.edu