Guidance on Research Registry and Research Repository Studies

This guidance document is intended to assist North Texas Regional (NTR) IRB investigators with crafting the appropriate information to include in the Protocol Synopsis of a project which is intended to be either a research registry (data bank) or a research repository (tissue/biospecimen bank). These types of studies involve collection and storage of information and/or biological specimens over time, and may involve the examination of one or more research questions (which will use information and/or samples collected in the registry or repository).

**When submitting a Research Registry or Research Repository study to the NTR IRB, researchers should complete the general Protocol Synopsis template, "Protocol Synopsis for Research Project Involving Human Subjects."*

INTRODUCTION:

Databases, registries (data banks), and repositories (tissue/biosample banks) all involve the collection and storage of information and/or biological specimens over time. Some registry/repositories serve diagnostic or clinical purposes, while others are created solely for research purposes. Many may serve more than one purpose. Rapid advances in clinical procedures have allowed registries and repositories to serve as tremendous resources for investigators. Specifically, registries and repositories allow for questions to be addressed that extend beyond those envisioned at the time of their creation.

Research uses for data and/or biospecimens that are stored in registries and repositories are governed by the Common Rule (aka, HHS Regulations for the Protection of Human Subjects in Research, at 45 CFR 46) and the HIPAA Privacy Rule (45 CRF 160 & 164), as well as by the NTR IRB policies and procedures and your local institutional policies/procedures. Specific requirements depend upon how and why the information or specimens are collected, stored, used, and shared. The requirements for and extent of IRB oversight depends on the whether or not the data/biospecimens include or are linked to individually identifiable health information and the terms of the informed consent under which the data/biospecimens were originally collected.

DEFINITIONS FOR THE PURPOSE OF THIS GUIDANCE DOCUMENT:

1) **Database:** A collection of information elements (i.e., data) arranged for ease and speed of search and retrieval.

   Examples of databases may include:

   - A set of observations (i.e., data) resulting from a research study (may include data collected from a clinical trial)
   - An electronic file containing patients' records

- A collection of diagnosis, treatment, and follow-up information for a hospital's oncology patients
- A file of outcomes information complied for quality assurance activities
- A list of names, diagnosis, and contact information developed and maintained to identify prospective research subjects
- A collection of medical information intended for use in future research studies

2) **Registry (Data Bank):** A collection of information elements or databases whose organizers:

- Receive information from multiple sources;
- Maintain the information over time;
- Control access to the information;
- Permit multiple individuals to use the information for a variety of purposes which may evolve over time;
- May (often) contain codes that link information and specimens to the donor's identify. When a key to the code is retained, it may be maintained either by the registry or by the provider of the data.

Registries may be publicly accessible or private. Examples of a few well-known registries and data banks include:

- Centers for Disease Control & Prevention (CDC)
- State Cancer Registries
- The National Library of Medicine Hazardous Substances Data Bank (HSDB)
- The National Practitioner Data Bank
- The US Census 2000 Data Bank

3) **Repository (Tissue/Biosample Bank):** A collection of biological specimens (biospecimens) whose organizers receive specimens from multiple sources. Activities of a Repository include:

- Maintaining the specimens over time.
- Controlling access to the biospecimens.
- Permitting multiple individuals to use the biospecimens for a variety of purposes which may evolve over time.

- Usually including phenotypic data (demographic and/or medical information) about the individuals from whom the specimens were obtained.
- When they do contain phenotype data, the repository is both a registry and a biospecimen repository.
- May (often) maintain codes that link the information and specimens to their donor's identify. The key to the code may be maintained by either the registry or by the provider of the data/biospecimen.

Examples of a few well-known repositories include:

- The National Human Radiobiology Tissue Repository
- The National Institute of General Medical Sciences (NIGMS)
- Human Genetic Cell Repository

SETTING UP A RESEARCH REGISTRY PROJECT:

Research databases may be maintained after the completion of a study. Additional questions may arise in the future that can be addressed using the same dataset. If there is an intent to set up a registry, then the informed consent document should include language for subjects to be aware of the storage of their data for future research purposes (and should give them the option to refuse this). The subject's decision as indicated on the consent/authorization must be respected and tracked. IRB oversight is required for each new research protocol that uses identifiable or re-identifiable information contained in the database.

Before proposing the establishment of a data registry for research, an investigator must consider whether the information they plan to collect would be readily available from an already established research registry (either one that already exists within the institution, or is accessible to the investigator by another means). Please consider and be mindful of the fact that the establishment of multiple independent registries collecting duplicate material increases the risk of tracking errors due to variability in practices and creates confusion on behalf of participants (and in turn, may also create confusion among investigators).

An additional note that the collection, storage, and distribution of personal identifying information (such as the 18 HIPAA Identifiers) for research purposes is subject to IRB review and human subjects research regulations. The IRB is charged with reviewing protocols for obtaining, storing and sharing information, verifying informed consent (when applicable), and protecting privacy and confidentiality. Since there is extensive variation in how registries operate, the IRB submission should include sufficient information regarding the scientific goals, functions, and operational procedures. Before establishing a registry, please review the section below on the information that will need to be included in your Protocol Synopsis (as well as consent document, if applicable).

Elements to include in the NTR IRB Submission of a Registry Protocol:

- The Purpose of the registry (include in the “Specific Aims” section of the Protocol Synopsis);

- The Background of why the creation of this registry is needed (include in the "Background and Significance" section);
- Include the following information in the "Methods and Procedures" section:
    - Scope of the data set, patient outcomes, target population, etc.
    - Data procurement - whether data will be extracted from a specific source (e.g., electronic medical record) or if data will be obtained through interaction with participants;
    - General description of the data/information to be collected for research purposes;
    - Immediate and/or future secondary use (may be unspecified);
    - Diagnosis or conditions of study (e.g., specific disease area or broad unspecified use);
    - How personal identifying information will be shared and procedures for coding, de-identification, encryption, data-use agreements, etc.;
    - If applicable, the consent process (who will obtain, documentation, place, time allotted, etc.) as well as the parental permission process and the assent process (if the study involves participants under the age of 18);
        - Additionally, if the study involves minors (participants under the age of 18 at the time of initial assent), the process of re-consent of these research participants who turn 18 while the registry is active also needs to be described.
    - Tracking participant choices where options are provided;
- Include the following information in the "Data Storage and Confidentiality" section:
    - Management (including levels of access) and physical storage of data (sample Data Storage and Security language can be found here: https://www.unthsc.edu/north-texas-regional-irb/wp-content/uploads/sites/61/Guidance-and-Procedures-for-Investigators-Data-Storage-and-Security-Final-October-2021.docx);
    - Length of time personal identifying information will be kept (indefinitely, end of research protocol);
- Include the following information in the "Human Subjects" section:
    - Whom research data will be collected from (e.g., minors, adults, healthy subjects, patients);
    - If applicable, the ability and procedure for locating/contacting participants (re-consent, incidental findings);
    - Approximate number of subjects to be included in the registry;
- Include the following information in the "Risk/Benefit Assessment" section:
    - Risk associated with a breach of confidentiality including impact on privacy, stigmatization, etc.;
    - Potential benefits to society, science, etc., from creation of this research registry, and how these benefits outweigh the potential risks.

Other Notes for NTR IRB Submission of a Registry Protocol:

- If Informed Consent and HIPAA Authorization *will not be sought* for the purposes of the registry protocol (e.g., participant data are being collected from review of medical records, and it would be impracticable to obtain informed consent and HIPAA authorization), a reminder to please

submit a completed Waiver of Informed Consent and a completed Waiver of HIPAA Authorization.

- If Informed Consent and HIPAA Authorization ***will*** *be sought* in order to include participants on the registry (e.g., there will be interaction with participants), please submit appropriate consent and HIPAA documents. *Example documents can be found on the NTR IRB Website:
    - *Consent Form Template (General)*
    - *HIPAA Authorization Template (UNTHSC Version)*
    - *HIPAA Authorization Template (JPS Version)*
- Please include a copy/sample of the data elements/variables to be collected as part of this registry (this can be done in whatever format is easiest for the investigator, such as an Excel sheet, Word document, etc.).
- If data elements are being received from, or transferred to, another institution/organization, please ensure that this information is appropriately outlined in the Protocol Synopsis, and also that any/all necessary agreements, etc., are in place (such as a Data Use Agreement).

SETTING UP A RESEARCH REPOSITORY PROJECT:

Biospecimens may be collected to achieve one or more of the objectives of a single study with disposal of the leftover materials at the end of the study. A repository is created if the leftover materials are stored for future use. If there is an intent to set up a repository, then the informed consent document should include language for subjects to opt-in or opt-out of storage of their data for future research purposes, a statement that the subject's biospecimens may be used for commercial profit and whether the subject will or will not share in this commercial profit, and whether the research will or might include whole genome sequencing. The subject's decision as indicated on the consent/authorization must be respected and tracked.

Specimens may also be collected specifically for the purposes of future research. IRB submission is required for each new research protocol that creates a repository.

The NTR IRB is concerned with the following three types of repositories:

1) Repositories of samples collected prospectively for a research study or possible future research study;
2) Repositories of clinical samples that were collected during medical procedures, to be used in the future for diagnostic and/or predictive purposes; and
3) Banks of clinical samples for which there is excess tissue/material (beyond which is needed for future medical diagnosis) that may be accessed for research purposes.

Repositories that include samples that will be used in research studies should have an appropriate IRB approved protocol in place to regulate the collection, storage, and distribution of samples. IRB review and oversight will be required for all such repositories that reside at an NTR IRB affiliated institution (see additional detail below).

Banks of clinical samples not intended for research may not be subject to federal regulations or IRB review unless the repository is federally funded or if IRB review is required by the institution. Institutions

typically do not require IRB approval for banks of clinical specimens that are not intended for research purposes. However, HIPAA regulations will apply to the clinical samples stored in these repositories. In all cases, it is important to avoid the regulatory problem created by collecting samples in a clinical care enterprise that are actually intended for research purposes. If there is any expectation that samples may be used for research purposes in the future, it is best to establish that assumption at the beginning, and create a research specimen repository in compliance with federal regulations. Consult with the NTR IRB staff for guidance on this topic.

Establishing Repositories for future research use:

Investigators who wish to create a repository of human samples that will be used for future research purposes will need to obtain IRB approval prior to establishing the repository. These Principal Investigators become "Repository Controllers" who establish and manage the collection, storage, access, and distribution of repository specimens. The IRB will review the operating procedures of the repository, including who will have access to the samples, coding of samples, and the process to ensure that future research projects are not conducted without prior IRB approval. Additionally, the IRB will consider ownership of the samples, privacy and confidentiality, process for withdrawing samples from the repository, plans for the transfer of samples to internal and external investigators, and oversight of future research involving the banked samples. The investigator may plan to use the banked samples for a variety of purposes. If so, this should be clearly described in the Protocol Synopsis. Please see below for guidance on the information to include in the Protocol Synopsis for a Repository.

Elements to include in the NTR IRB Submission of a Research Repository Protocol:

- The Purpose of the repository (include in the "Specific Aims" section of the Protocol Synopsis);
- The Background of why the creation of this repository is needed (include in the "Background and Significance" section);
- Include the following information in the "Methods and Procedures" section:
    - Scope of the biosamples (and associated data, if applicable) to be collected, patient outcomes, target population, etc.
    - Procurement - whether biosamples (and associated data, if applicable) will be obtained from a specific source or if the samples will be obtained through interaction with participants;
    - General description of the biosamples (and associated data, if applicable) to be collected for research purposes;
    - Immediate and/or future secondary use (may be unspecified);
    - Diagnosis or conditions of study (e.g., specific disease area or broad unspecified use);
    - How personal identifying information will be shared and procedures for coding, de-identification, encryption, material transfer agreements, data-use agreements, etc.;
    - If applicable, the consent process (who will obtain, documentation, place, time allotted) as well as the parental permission process and the assent process (if the study involves participants under the age of 18);

        - Additionally, if the study involves minors (participants under the age of 18 at the time of initial assent), the process of re-consent of these research participants who turn 18 while the repository is active also needs to be described.
    - Tracking participant choices where options are provided;
- Include the following information in the Data Storage and Confidentiality section:
    - Management and physical storage of biosamples, and description of the levels of access;
    - Length of time any personal identifying information associated with the biosamples will be kept (indefinitely, end of research protocol);
    - For any data to be collected (with the biosmaples), please refer to our example Data Storage and Security language, which can be found here (a reminder/note that this language would *only* apply to the data/information collected, and *not* the biosamples): https://www.unthsc.edu/north-texas-regional-irb/wp-content/uploads/sites/61/Guidance-and-Procedures-for-Investigators-Data-Storage-and-Security-Final-October-2021.docx)
- Include the following information in the "Human Subjects" section:
    - Whom research samples and data (if applicable) will be collected from (e.g., minors, adults, healthy subjects, patients);
    - If applicable, the ability and procedure for locating/contacting participants (re-consent, incidental findings);
    - Approximate amount of samples/subject information to be kept within the repository;
- Include the following information in the "Risk/Benefit Assessment" section:
    - Risk associated with a breach of confidentiality or accidental loss/destruction of samples including impact on privacy, stigmatization, etc.;
    - Potential benefits to society, science, etc., from creation of this research repository, and how these benefits outweigh the potential risks.
    - If human genetic research is anticipated, please include possible consequences of genetic testing (e.g., insurance risks, paternity determinations, etc.) and related confidentiality risks.

Consent Process for a Research Repository:

Samples that will be prospectively obtained from subjects will require the informed consent of the subject before they can be placed in the repository. This consent should include the core elements described in the federal regulations at 45 CFR 46.116 (b). A reminder that for samples in which identifiable information is linked, appropriate HIPAA authorization would also need to be obtained.

Additionally, it may be appropriate to provide an "opt in" and "opt out" option in the consent form that allows subjects to determine which types of entities/individuals may use their tissue for research (i.e. nonprofit, commercial, specific researcher) and for what purpose the samples may be used (i.e. cancer research, cardiovascular health research, gynecological research, etc.). This is especially important for banked samples that will be used for a variety of research purposes. Existing samples that an investigator wishes to place in a repository may qualify for a Waiver of Informed Consent or they may

require the consent of the subject. Establishing a repository using donated and/or purchased samples is also discussed in greater detail later below.

Other Notes for NTR IRB Submission of a Registry Protocol:

It is important for investigators to remember that after the repository has been established, each individual research project that will utilize samples from the repository will require an individual IRB review and approval (the level of IRB review will be determined at the time of review and will be dependent on the activities being conducted). However, the individual(s) responsible for the oversight of the repository will need to ensure that access to the banked samples is only granted with IRB approval. Further, Repository Controllers (see above) should also be included as key personnel on the protocol to verify access and authorization for repository specimens and their associated data.

Additionally, in some circumstances, the IRB may require that a person whose sample is in the repository provide additional consent (i.e. be "re-consented") to allow researchers to use their sample for their research. An example would be research that involves HIV testing of stored samples. This may involve re-contacting subjects to obtain their consent. Including an "opt in" and "opt out" clause in the consent form may reduce the need to re-contact subjects. Guidance on re-contacting subjects can be found on the North Texas Regional IRB website under at the following link: https://www.unthsc.edu/north-texas-regional-irb/wp-content/uploads/sites/61/Re-Consenting-Subjects_Revised-February-2018-Final.pdf

Research using Existing Donated or Purchased Samples

- ***Creating a Repository using Donated or Purchased Samples***

Research studies at NTR IRB affiliated institutions may involve the collection and analysis of existing samples and related health information donated by, or purchased from, an outside entity or individual. The investigator may wish to create a repository that includes these donated samples and related health information. The related health information may be extensive in some situations (for example, an entire medical chart), and include a great deal of Protected Health Information (PHI). Investigators will be required to describe, in the protocol synopsis, to what extent the samples and related data will be de-identified before arriving at their institution. Additionally, investigators should also describe, again, in the protocol synopsis, the process in which identifiers will be "stripped" from the medical data. If the breadth of health information is extensive, the IRB will require that all identifiers be removed before the samples arrive at the research site.

- ***Ensuring Donated or Purchased Samples are Legally and Ethically Obtained***

It is important that investigators ensure that the samples received from the outside entity were and continue to be legally and ethically obtained. Documentation describing how the samples were obtained should be submitted with the IRB application. This may be demonstrated by obtaining a copy of IRB

approval for the collection of samples from the outside entity. Please Note: A copy of the outside entity's IRB approval should be submitted to North Texas Regional IRB with the IRB submission.

- ***Ownership of Samples***

Investigators should describe, in the protocol synopsis, who will own the samples and related data after they arrive at the research site, and indicate if the outside entity/individual will retain any ownership or access to the samples after they are transferred to the research site.

Transfer of Samples and Related Data to other NTR IRB Affiliated Researchers

An investigator serving as the Repository Controller may wish to allow other researchers (whose institutions are affiliated with the NTR IRB) to use samples stored in a repository for future research purposes. In this case, a section titled "Transfer of Specimens and Data to Other Researchers" should be included in the protocol synopsis. This section should describe the process for how specimens and data will be transferred to other investigators at the NTR IRB affiliated institution. As mentioned earlier, an entirely new IRB submission will be required before an investigator can access these samples or data for their individual research project.

Transfer of Samples and Related Data to Outside Researchers

An investigator may wish to transfer samples to an outside researcher for several reasons. The outside researcher may be involved in the analysis of the samples and related data for a current research project the investigator is conducting, or for a collaborative research project conducted by the NTR IRB affiliate, and an outside entity. Additionally, an investigator may wish to allow an outside researcher to use samples stored in a repository for a future research project that does not involve the NTR IRB affiliate.

In all cases, an appropriate set of procedures will need to be in place to protect the subjects' confidentiality during this transfer. Investigators are encouraged to consider such future arrangements and to establish these procedures within the initial protocol application. However, there may be situations when it is necessary to modify an existing IRB approved protocol to include this option. The protocol should include a section titled "Transfer of Specimens and Data to External (Non-NTR IRB) Researchers" that describes a detailed plan for the transfer of samples to outside researchers. To protect subject confidentiality, the protocol should describe how the samples and data will be labeled when they are transferred, and the process for "stripping" all identifiers from the data before they leave the institution. The protocol should also list the outside researchers by name, and describe who will own the samples after they are received by the outside investigator if this information is available. The IRB understands that an investigator may not be able to name a specific outside researcher during the initial IRB application process or request for modification to the protocol, however would like to include this option should there be a future need to transfer samples.

Once the Principal Investigator (PI) has identified the outside entity who will receive and analyze the specimens, that PI is required to submit a request for approval to the North Texas Regional IRB. This memorandum should:

- Name the outside entity,
- Describe what the outside entity will do with the samples,
- Describe who will own the samples after they are received by the outside entity, and
- How the samples and accompanying data will be securely maintained.

The investigator should not send the samples and/or related data to the outside entity until they have received notice of approval in writing from the North Texas Regional IRB.

Subjects should be advised during the initial informed consent process that their samples and/or related data may be sent to an outside entity that is approved by the IRB for research purposes. Only subjects who consent to the transfer of specimens should have their samples and/or related data sent to outside researchers. In some cases, it may be necessary to re-contact and re-consent subjects who were advised that their samples and/or related data would not be transferred to outside researchers during the initial informed consent process. Investigators are encouraged to contact the IRB staff for guidance in this area. *This will not be applicable for studies that qualify for a Waiver of Informed Consent.

Research Using Samples from Deceased Persons

The definition of *human subjects research*, per the Federal Regulations (at 45 CFR 46), does **not** include deceased persons. Therefore, the use of samples obtained during an autopsy or the use of samples originally collected from a living individual who is now deceased is not considered research with human subjects. However, in most cases, other federal and state regulations may apply including HIPAA regulations. Investigators are encouraged to contact the IRB staff prior to initiating this type of research to ensure that appropriate HIPAA and/or IRB compliance is followed.

***The NTR IRB thanks the Children's Hospital of Philadelphia Institutional Review Board, the University of Miami Human Subjects Research Office, as well as the University of Kentucky Office of Research Integrity for allowing guidance from each of their institutions to be used in this document.**