Yale University
Human Research Protection Program (HRPP)
P.O. Box 208327
New Haven, CT 06520-8327
Phone: 203-785-4688

Guidance: Yale Repositories		
Document #:	**Revision #:**	**Effective Date:** 11/01/2025

1. PURPOSE

The purpose of this guidance is to ensure that the human data/biospecimens banked in repositories to be used for research purposes are responsibly obtained, stored, and distributed, and to protect the rights of individuals from whom data and biospecimens were obtained. This guidance outlines Yale University and Yale Human Research Protection Program (HRPP) requirements for the creation and maintenance of research repositories.

2. SCOPE

This guidance applies to human subjects research repositories established by Yale University or Yale New Haven Health System (YNHHS) faculty, students, and staff for the purpose of acquiring, storing and/or distributing human data/biospecimens for research purposes.

3. HRPP POSITION ON REPOSITORIES

3.1. **Ownership:** It is stipulated that research participants contribute data/biospecimens to Yale University, with the University retaining the ownership of these data/biospecimens. Yale University also retains ownership of any derivatives of biospecimens irrespective of custodianship unless otherwise stipulated by all parties to the research agreements for which biospecimens were collected. If an investigator places data/biospecimens into a Yale University repository and subsequently leaves Yale University or its affiliate institutions, the data/biospecimens remain the property of the University and may not be taken from the repository without permission from the Vice Provost of Research. This is consistent with Yale University Research Data & Material Policy, §6001.1.

3.2. **Human subject specimens and data may not be sold for profit:** Investigators are allowed to recoup the actual costs (e.g., collection, processing, storage and distribution) incurred in preparing and transporting repository samples, but in no case should a profit be generated. Repository cost recovery must be in accordance with other relevant University guidelines and policies.

3.3. **Roles and Responsibilities:** All repositories, regardless of type, must designate at minimum, two key roles: Principal Investigator and Repository Guardian. The Principal Investigator should also designate an Honest Broker for repositories approved to contain only anonymized, deidentified, and/or coded material. The Principal Investigator may choose to designate an Honest Broker in other circumstances as well. If the repository requires submission to and approval of an IRB, the IRB application should delineate the responsibilities of each role.

3.3.1. **Principal Investigator (PI)** is ultimately responsible for all aspects of the repository, including actions and responsibilities delegated to other study team members, such as the Repository Guardian. The PI must ensure that the repository is managed as described in the IRB approved protocol (including maintenance of coding, anonymization, and/or deidentification of data/biospecimens); submit any changes to the repository prior to implementation; and not interfere with or influence the Honest Broker's performance of required functions; and closure of the repository or transfer of the repository if the PI leaves.

3.3.2. **Repository Guardian** has primary control of data and biospecimens and maintenance of the repository. This person may be the PI or a study team member delegated by the PI. However, the PI retains ultimate responsibility for oversight of the repository. The Repository Guardian may be responsible for the following:

- Ensuring that data/biospecimens are received and released according to Yale University policy and the IRB approved repository protocol;
- Executing a Repository Sharing Agreement prior to the release of data/biospecimens for research purposes;
- Ensuring the security and confidentiality of stored data/biospecimens;
- Ensuring the security and confidentiality of data/biospecimens during transfer/distribution;
- Tracking acquisitions and releases of data/biospecimens;
- Maintaining methods for identifying data/biospecimens for which consent and/or authorization have been withdrawn and ensuring no future research use;
- Identifying data/biospecimens that have limitations on future uses and ensuring future uses do not violate those limits.

3.3.3. **Honest Broker** does not work with or directly report to the PI or other key personnel and is not involved in the repository or resulting research. This person can generate or receive data/biospecimens and then remove identifiers to create anonymized, deidentified, and/or coded or anonymized material. The Honest Broker must not share identifiers or the key to codes with the PI or study team.

3.4. **Data and Systems Security:** Yale protects IT systems based on risk using Minimum Security Standards (MSS). These standards are the baseline security requirements for protecting Yale IT systems based on the risk they carry. There are two key elements to consider when thinking about the risk: 1) data classification identifying the sensitivity of the data; and 2) external obligations to determine whether the data is subject to third-party controls (e.g., HIPAA, FERPA, etc.). Investigators must evaluate the risk level of repository data and either use an existing service that matches the classification of the data or build or purchase a new system that meets MSS appropriate for the data's risk level.

3.4.1. **Business Associates (BA):** Repository data that is subject to HIPAA may require execution of a Business Associates Agreement (BAA) if an entity or person will provide a function involving the use or disclosure of protected health information (PHI) on behalf of a Yale Covered Entity (CE) component or provides certain specified services where the provision of the service involves the disclosure of PHI for a CE. Before engaging with a Business Associate, investigators must contact HIPAA@yale.edu to determine whether a sufficient BAA is already signed and on file. If an agreement does not exist yet, the BAA must be executed with the Business Office and Yale HIPAA Privacy Office.

3.5. **Federal Funding Sharing Requirements:** Federal agencies have varying requirements for data management plans, and almost all require a data sharing policy. Repositories receiving federal funding must check the agencies' requirements to ensure data are accessible, interoperable and reusable, and adhere to privacy and security requirements according to the agency's specific requirements. Plans for data management and sharing must be described in the IRB-approved repository protocol.

3.6. **Attribution:** To ensure transparency and recognition of the resources and efforts that support research endeavors, investigators who use human biospecimens or data from a Yale University general use research repository must acknowledge the University repository as the source of human biospecimens or data in all publications, proposals or presentations that result from the use of the human biospecimens. Additionally, recipients of human biospecimens or data from a research-specific repository under the guardianship of another individual must follow the University's guidelines on attribution and authorship.

4. OBTAINING IRB APPROVAL FOR A REPOSITORY

4.1. IRB approval must be obtained to establish a repository. Repository activities may not begin until IRB approval has been received.

4.2. IRB approval or exemption determination must be obtained to initiate research that utilizes identifiable data/biospecimens from a repository.

4.3. Repositories must be submitted as separate and distinct protocols with funding that supports the management of the repositories (and not funding that provides support for analysis of research using data/biospecimens from the repository). Similarly, research using identifiable data/biospecimens from a repository must also be submitted as separate and distinct protocols, including research conducted by repository investigators and/or co-investigators using their own banked material for research purposes or aims that are not already described in the repository protocol. This guarantees that each protocol undergoes an independent review process, addressing specific aims, methodologies, and unique ethical considerations, and ensures compliance with ethical standards and regulatory requirements.

4.4. A protocol for creating and maintaining a repository must include all typical elements found in human subjects research protocols, such as the purpose, consent process, and data safety monitoring. Additionally, the protocol should provide detailed descriptions specific to the repository's functions, such as the distribution of samples for secondary data research. Researchers interested in developing a data or tissue repository can access a ***Repository Protocol*** template. A repository research protocol should, at a minimum, include the following key components:

- The purpose of the repository, including why creation of the repository is needed (include references to support this conclusion);
- Maintain the identities of key repository personnel (e.g., PI, Repository Guardian and Honest Broker, if one will be used);
- The immediate and/or future secondary uses (may be unspecified) of the data/biospecimens;
- Whom the data/biospecimens will be collected from and the approximate amount of data/biospecimens to be kept within the repository; include the scope or type of data/biospecimens to be collected, including frequency, volume, size, etc.;
- How the data/biospecimens will be procured (e.g., retrospectively, prospectively, or both); distinguish between standard of care clinical practices and dedicated research activities.

4.4.1. **Collection of data/specimens.** The repository protocol must describe whether the repository involves data/biospecimens collection at a single time point or if longitudinal collections will or may be performed. If data/biospecimens will be collected longitudinally, describe the frequency/collection intervals. If biospecimens will be collected longitudinally, describe the maximum amount of material (e.g., blood, biopsy size, etc.) to be collected at each timepoint and in total (when possible) Specify whether the biospecimens are collected for research purposes only, or if discards will be taken from standard of care clinical procedures. If applicable, describe the plan for certifying documentation of local IRB approval to the Yale University IRB for each site contributing data/biospecimens to the Yale University repository.

4.4.2. **Risks and benefits of the repository**. Give special consideration to risks associated with a breach of confidentiality or accidental loss/destruction of samples including impact on privacy, stigmatization, etc. If human genetic research is anticipated, include possible consequences of genetic testing, whether incidental findings will be disclosed to the participant, and related confidentiality risks. When creating a repository specific to cultural groups such as native tribes, it is essential to consider and describe possible negative impacts that can affect the entire group. These impacts could include:

- Stigmatization: Data pointing to particular health issues could lead to negative stereotypes about the tribe.

- Loss of Autonomy: By surrendering control over their data, the tribe might lose its ability to manage how this information is used or shared, which can be contrary to their cultural or social norms.
- Cultural Infringement: The data might include culturally sensitive information that, if shared or misinterpreted, could infringe on the tribe's customs, practices, or belief systems.
- Economic Exploitation: The data could be used in ways that economically benefit others without providing any benefit, or potentially causing harm, to the tribe itself.

Describe how these risks will be mitigated. For example, researchers should engage with the group from the outset, obtain informed consent from the leaders and members of the group, and ensure cultural sensitivity in data handling. Such measures include collaborating with tribal leaders throughout the research process, respecting data sovereignty, and adhering to ethical standards to safeguard the group's interests and rights.

4.4.3. **Participant Identification.** Include a clear description of how potential participants will be identified. This applies for repositories involving prospective collection of data/biospecimens performed by the study team and situations where investigators will obtain existent material from other sources (e.g., clinical discards; donations and/or purchases from external repositories, marketing firms, etc.).

4.4.4. **Age of Majority Considerations:** If the repository involves minors, a process for obtaining consent from participants when they reach the age of majority while their participation in the repository is active must also be described. If the consent, parental permission, and/or assent process allows participants to opt in or out of specific activities or future uses, describe how participant choices will be tracked.

4.4.5. **Dual Enrollment of Participants:** The repository protocol should describe any measures that will be taken to identify potential participants who have enrolled in multiple repositories or have participated in other research at Yale. At minimum, potential participants should be asked about any prior or current participation in Yale research and notified of potential conflicts, including risks of dual enrollment in multiple studies (such as exceeding permissible research blood volumes [see HRPP Supplemental Guidance Manual – Reference Guide: Blood Limit Guidelines for Adult and Pediatric Research]) and the possibility of being ineligible for the repository due to simultaneous participation in multiple studies.

4.4.6. **Data/Biospecimen management and confidentiality**. Ensure the application adequately addresses:

- Physical storage of data/biospecimens, including location/housing, and levels of access. Methods for preventing unauthorized access to repository materials (such as computer security and physical storage security measures) should be described. Access to identifiable data/biospecimens must be restricted to a limited number of repository staff.
- Level of identifiability of the data/biospecimens in the repository. If applicable, describe the methods for de-identification/anonymization or coding of data/biospecimens. If coded, describe the process for limiting access to the key.
- Procedures for intake and release data/biospecimens, including preparation and execution of proper institutional agreements (e.g., Data Use Agreement, Material Transfer Agreement, Sharing Agreement, etc.); consent obligations; and protecting privacy and confidentiality during transfer (e.g., encryption, other types of restricted access).
- Length of time any personally identifying information associated with data/biospecimens will be retained (indefinitely, end of protocol, other contractual obligations, etc.).
- How/whether participants will be able to withdraw data/biospecimens from the repository and/or receive research results.

4.4.7. **Considerations for enrollment of non-English speaking individuals:** When considering enrollment of non-English participants in a repository protocol, researchers must prioritize inclusivity and ethical considerations. The research protocol must describe measures to ensure that non-English speaking participants will be provided with clear, comprehensible information about the study, ensuring they fully understand the purpose, procedures, risks, and benefits. This must involve translating consent forms (documentation of consent using short form is highly discouraged), informational materials, and surveys into the participants' native languages. Additionally, having bilingual staff or translators available can facilitate effective communication and support. Researchers should also be sensitive to cultural differences that may impact participation and work to build trust within diverse communities. Ensuring that language barriers do not hinder participation is crucial for achieving a representative sample and enhancing the generalizability of the research findings.

4.5. **Consent from participants**. Research repositories that involve prospective collection of data/biospecimens require either documented consent from participants (including parental permission and/or assent of minors) for the storage and future research use of their data/biospecimens or a waiver of consent by the IRB.

4.5.1. **Informed Consent Document:** The consent form must contain all the basic elements for consent required by federal regulations and must include a clear statement that the participants are giving permission for their data/biospecimens to be stored in the repository to be used in future research studies. The statement will dictate the limits on future uses. In addition, the Consent Forms and the informed consent discussion should describe all the following to the potential participant:

- A clear description of the repository, including how data/biospecimens will be obtained; what types of research the data/biospecimens will be used to investigate; what information may be released (e.g., level of identifiability); and a description of the procedures for protecting confidentiality, including storage and sharing of data/biospecimens;
- What types of investigators may request data/biospecimens from the repository (e.g., Yale University and Yale New Haven Health System investigators, external institutions, industry, government, etc.);
- How requests for release of data/biospecimens will be reviewed to ensure the research is consistent with the aims of the repository (unless it is a general use repository or otherwise stipulated) and the potential participant's consent and authorization;
- A statement regarding whether and how participants may, in the future, request that their data/biospecimens be destroyed or that all personal identifiers be removed from their data/biospecimens;
- Other information, such as the length of time the data/biospecimens will be stored; participants' access to information learned from research resulting from use of repository materials; and secondary uses of data/biospecimens (if appropriate);
- If potential participants will have an opportunity to opt in or out of any elements of the repository (e.g., participate in the main study but not the repository component or substudies; provision of specific biospecimen types; place limitations on future uses), the form must be adequately designed so that the potential participant marks the exact elements of the repository they agree to participate in. Investigators must track responses and ensure that data/biospecimens are appropriately labeled and/or access is appropriately restricted.

4.5.2. **Waiver of Consent:** A waiver of consent may be granted by the IRB if all requirements for a waiver are met. Data/biospecimens that are collected and stored in a repository under a waiver of consent may have limitations on future use and other restrictions.

4.6. **Authorization to Use and Disclose Protected Health Information (PHI):** In addition to consent, research repositories that are subject to HIPAA regulations require either documented HIPAA Authorization from participants for the storage and future research use of their data or a waiver of authorization by the IRB serving as Privacy Board.

4.6.1. **HIPAA Authorization:** The HIPAA authorization form must contain all the core elements and required statements. Authorization language can be incorporated into a consent form as a Compound Authorization and Consent document or can be obtained using a stand-alone HIPAA Research Authorization form.

4.6.2. **Waiver of Authorization:** A waiver of authorization may be granted by the IRB serving as Privacy Board if all requirements for a waiver are met. Data that are collected and stored in a repository under a waiver of authorization may have more limitations on future uses.

4.7. **Distribution of Data and Samples:** The protocol must describe conditions under which the data and samples from the repository will be distributed to other researchers, ensuring the distribution plan aligns with the overarching goals of the repository. Specifically, the plan should include the following elements, however, the detailed descriptions of each of these can be included in the repository SOPs, which do not require submission to the IRB:

- Eligibility Criteria for Distribution: Define the criteria for determining which researchers or organizations are eligible to receive data and/or samples. Include considerations such as:
    - Ethical approval by an Institutional Review Board (IRB) – the distribution plan can distinguish between requirements for sharing deidentified data or samples for a specific use (no IRB determination needed) and sharing identifiable data/samples or deidentified data when a merger with a different dataset is planned (IRB exemption or approval needed);
    - Researcher's affiliation with a recognized research/academic institution;
    - The intended use of the data and/or samples and its alignment with the original consent provided.
- Request Process: Detail the procedure for requesting access to data and/or samples. This could include:
    - Submission of a formal request or application.
    - Requirements for supporting documentation, such as research proposals or IRB approvals.
    - Contact information for the repository's guardian or governance body.
- Review and Approval Process: Outline the steps involved in reviewing and approving distribution requests, including:
    - Description of the party responsible for evaluating requests. The PI and the Repository Guardian may determine that a governance body should be formed to review certain requests for access to identifiable data. For example, when access is requested by external institutions or commercial entities; or when the samples raise ethical issues, such as those involving vulnerable populations or controversial research areas, a governance body can provide a thorough review to ensure alignment with ethical standards and societal values.
    - The criteria and processes used to assess requests, including relevance, scientific merit, and ethical considerations.
    - Expected timelines for the review and approval process.
- Data and Sample Handling and Transfer: Specify the protocols for handling and transferring data and/or samples to ensure security and integrity, such as:
    - Packaging and shipping procedures.
    - Storage conditions during transit.

    - Digital data transfer methods, including encryption and secure file sharing platforms.
- Recipient Obligations: Clearly state the obligations and responsibilities of the recipients, including:
    - Compliance with ethical guidelines and regulations.
    - Proper acknowledgment of the repository in publications and presentations.
    - Submission of progress reports or final outcomes of the research conducted using the distributed data and/or samples.
    - When applicable, provide information about Certificate of Confidentiality requirements.
- Confidentiality and Privacy Protections: Detail measures to protect the confidentiality and privacy of the data and/or samples being distributed, including:
    - Anonymization, pseudonymization (for data subject to GDPR), or de-identification processes. Note that DICOM images (e.g., x-rays, MRIs, etc.) are very difficult to deidentify as the metadata file contains a date, which is one of the HIPAA identifiers. As such, you should treat these images are identifiable data subject to HIPAA Privacy Rule. Review Sharing Data for Research Purposes guidance for additional information about the deidentification processes.
    - Requirements for recipients to maintain confidentiality and security of the received data.
- Monitoring and Compliance: Describe monitoring practices to ensure compliance with the distribution plan, such as:
    - Periodic audits or reviews of recipient usage.
    - Procedures for addressing breaches or non-compliance.

5. POST IRB-APPROVAL REPOSITORY ACTIVITIES

5.1. Collecting, Storing, and Managing Repository Data/Biospecimens

5.1.1. **Prospective Collection:** All activities involving the prospective collection of data/biospecimens through interaction or intervention with living individuals must be conducted under the terms of an IRB-approved repository research protocol. The prospective collection (either identifiable or deidentified) must comply with all applicable laws, regulations, and institutional policies. All signed consent forms must be kept in the regulatory binder. The retention period of the consent forms and all other documents is the same as other research projects.

5.1.2. **Obtaining Material from Outside Entities:** Investigators may collect data/biospecimens from outside entities or individuals (e.g., via donation or purchase) following execution of a Submittal Agreement. In such cases, investigators should ensure that the data/biospecimens were and continue to be legally and ethically obtained by obtaining documentation from the outside entities describing how the data/biospecimens were originally obtained. IRB approval letters for the collection of data/biospecimens from the outside entities and copies of the template (unsigned) consent/authorization form(s) allowing for donation of the samples to the repository should be maintained in the study regulatory binder. If the donation of the data or biospecimens to the

repository has not been part of the approved consent form, contact the IRB to discuss any available options. For example, if the consent form signed by the participant has not specifically preclude sharing and that activity otherwise meets criteria for a waiver, an IRB may issue a waiver of consent and HIPAA authorization, as appropriate, to include the data in the repository.

5.1.3. **Publicly Available Material:** Investigators may obtain material from public repositories for research. It is the investigator's responsibility to review the repository's terms of use, access restrictions, any licenses attached to the material, and attribution requirements to ensure the material can be put to the intended research use. Yale's Office of Sponsored Projects (OSP) and/or Procurement will handle necessary research agreements (e.g., purchase contracts, service agreements).

5.1.4. **Storage:** Data/biospecimens must be stored as described in the IRB-approved protocol without exception. Any changes in storage must be submitted to and approved by the IRB prior to implementation. In the event of a storage error, the IRB and/or Privacy Office (for HIPAA-regulated data) must be notified within required reporting timeframes.

5.1.5. **Updates/Management:** If data/biospecimens will be collected longitudinally, participant records must include documentation of each collection as well as re-consent if the study consent has been updated and re-consent was required by the IRB.

5.2. **Handling Requests to Distribute Data/Samples**

The investigators must establish a process for reviewing requests for and provisioning access to data/biospecimens in accordance with the IRB approved repository protocol. Details of the processes for handling requests should be described in the Repository SOPs. They should address the following elements:

5.2.1. **IRB Approval:** Separate IRB approval is required for each specific human subjects research activity that uses identifiable data/biospecimens from the repository. It is the recipient investigator's responsibility to obtain IRB approval and provide it to the Repository Guardian along with the request. Under no circumstances should a repository investigator retrieve or share identifiable data/biospecimens from a repository without first confirming the IRB approval status of the recipient investigator's research. The repository SOPS can describe specific situations when IRB approval/determination letter would not be required e.g., only when sharing data that has been deidentified by removing all of 18 HIPAA identifiers. However, due to limitations in identifying images, the Repository SOPs should state that IRB determination letter with documented HIPAA waiver for disclosure and access to PHI will be required.

5.2.2. **Verification of Request:** Before retrieving or sharing any data/biospecimens, repository investigators must first verify that the requested use is not contrary to any previously imposed limits, via law, previous consent, or other applicable limits. At minimum, investigators must review the participant's most recent signed consent and authorization form(s) to confirm that the new, proposed use aligns with the future uses described in the signed consent and authorization form. If the proposed use does not align with previously imposed limits, the investigators must decline to retrieve and share the data/biospecimens or work with the IRB to modify the study and obtain participant re-consent for the new, proposed use.

5.2.3. **Sharing Agreement:** Repository investigators must obtain a signed Repository Sharing Agreement from recipient investigators regardless of whether the data/biospecimens being shared are identifiable or deidentified. Repository investigators must save all copies of fully executed Repository Sharing Agreements in the research record and should be prepared to share copies with the HRPP on request.

5.2.4. **Sample Transfer:** If a recipient investigator requests biospecimens (samples) from the repository to a recipient external to Yale, investigators must work with the Office of Sponsored Projects to execute a Material Transfer Agreement (MTA). The Material Transfer Agreement should describe methods to ensure the stability and security of the biospecimens during their release and transfer to the recipient investigator, consistent with the IRB-approved protocol. If biospecimens are damaged, destroyed, or lost while in transit, the investigator must document the issue in the research record and notify the IRB either at continuing review or via a Reportable New Information (RNI) submission (if criteria for prompt reporting are met).

5.2.5. **Data Transfer:** If a recipient investigator requests data from the repository to a recipient external to Yale, investigators must work with the Office of Sponsored Projects to execute a Data Use Agreement, and the data transfer must be performed in a manner consistent with the IRB-approved protocol to ensure the data is kept secure and confidential. If data are misdirected or corrupted during transfer, the investigator must document the issue in the research record and notify the IRB either at continuing review or via a Reportable New Information (RNI) submission (if criteria for prompt reporting are met).

5.2.6. **Recipient Investigator Responsibilities:** Recipient investigators agree, via the Repository Usage/Sharing Agreement, Material Transfer Agreement, and/or Data Use Agreement, to:

5.2.6.1. Comply fully with all conditions described in the agreements;

5.2.6.2. Not to make any attempts to re-identify any coded or deidentified data/biospecimens; and

5.2.6.3. Promptly report to the investigator (or their repository guardian designee) any proposed changes in their research projects and any unanticipated problems involving risks to participants or others. If the recipient investigator changes the scope of their research, they are obligated to provide evidence of the IRB's approval of the modified research and update the Repository Usage/Sharing Agreement to describe the new use. Investigators must re-confirm that the modification is not contrary to previously imposed limits before the recipient investigator puts the data/biospecimens to the new use.

5.2.7. **Tracking Requirements:** All retrievals and sharing of repository data/biospecimens must be tracked and maintained in the research record under a tracking log. The Tracking log must include protocols under IRB review and non-human subjects research protocols. The tracking log must be made available to the HRPP if requested See Appendices 1 and 2 for required tracking log.

5.2.8. **Withdrawal and Destruction:** Investigators must maintain methods for identifying participants who have withdrawn consent and/or authorization to ensure their data/biospecimens will not be used in any future research. In general, investigators must destroy all data/biospecimens that remain in the repository unless participants allow use of deidentified data/biospecimens (and request destruction of identifiers only). Investigators are not expected to retrieve or destroy material that has already been shared with other researchers.

5.3. **Modifications:** Revisions to a research repository protocol must be approved by the IRB prior to implementation. If the proposed revisions are likely to affect the participants' willingness to continue allowing research use of their data or biospecimens, investigators must include a plan for notification of participants or obtaining re-consent from them. If the IRB requires re-consent and it cannot be obtained, participants' data/biospecimens must be stored, used, and shared as described in each participant's most current, signed consent form.

5.4. **Continuing Reviews:** Research repositories that are approved with the requirement for continuing review are required to provide a summary report to the IRB of all collections, releases, withdrawals, and destructions of data/biospecimens from the repository at each continuing review. The IRB may require summary reports from repositories that are exempt from continuing review on an ad hoc basis.

6. REPOSITORY CLOSURE

6.1. **Repository Closure:** When there is no intent to continue operating a repository or if all of the data/biospecimens are being transferred to another repository, the repository must be closed by submitting a Study Closure to Yale IRB via IRES IRB. The closure request must include a description of the disposition of the data/biospecimens, including transfers, donations, and destruction of all repository

material in a secure manner. If participants will be notified of the repository closure and/or disposition of the data/biospecimens, the communication (e.g., letter) must also be submitted to Yale IRB for review.

7. DEFINITIONS

7.1. **ANONYMOUS:** Data/biospecimens that do not include direct or indirect identifiers at any point in the research and cannot be retrieved by the investigator.

7.2. **ANONYMIZED/DEIDENTIFIED:** Data/biospecimens from which identifying information is removed. The remaining data cannot include any HIPAA identifiers or information that could have the potential for deductive disclosure.

7.3. **BROAD CONSENT:** A type of informed consent that allows a subject to agree to the storage, maintenance, and secondary research use of identifiable private information or identifiable biospecimens for future, unspecified research. As of June 2025, Yale University has not adopted the broad consent provisions and does not grant exemption determinations under categories 45 CFR 46.104(d)(7) (Storage or maintenance for secondary research for which broad consent is required) or 45 CFR 46.104(d)(8) (Secondary research for which broad consent is required).

7.4. **CODED:** Data/biospecimens labeled with a unique code (e.g., study ID). A separate link (key) is kept which connects this ID number to an identifier (e.g., name, medical record numbers, etc.). Data/biospecimens are usually coded if an investigator anticipates that they may need to gather additional data or verify data at more than one point over the life of the study. If a link exists, data are considered indirectly identifiable and not anonymous, anonymized or de-identified**.**

7.5. **DATA USE AGREEMENT:** A legally binding agreement required under the HIPAA Privacy Rule when a covered entity (e.g. hospital or health plan) discloses a limited data set to another party for research, public health, or heath care operations. The Data Use Agreement (DUA) outlines the permitted uses and disclosures of the limited data set and the recipient's obligations to protect the data.

7.6. **FUTURE USE:** Refers to the potential use of research data/biospecimens for purposes that may be related or distinct from the original study for which they were collected. The use of data/specimens for future research purposes should be specified in the consent form of the original study and may require additional approval or oversight.

7.7. **HUMAN SUBJECT:** A living individual about whom an investigator conducting research: (1) Obtains information or biospecimens through intervention or interaction with the individual and uses, studies, or analyzes the information or biospecimens or (2) Obtains, uses, studies, analyzes, or generates identifiable private information or identifiable biospecimens. (45 CFR 46.102(e)(1))

7.8. **HUMAN BIOSPECIMEN:** Encompasses a wide range of human specimen types, including but not limited to: sub-cellular components (such as DNA or RNA); cells or tissues from any part of the human body; organs (such as liver, kidney, heart, placenta, etc.); gametes (ova and sperm); embryos and fetal tissues; stem cells; bodily products (such as teeth, hair, nail clippings, sweat, urine, feces); blood and blood fractions (such as plasma, serum, buffy coat, red blood cells); and saliva and buccal cells.

7.9. **IDENTIFIABLE PRIVATE INFORMATION:** Information where the identity of the subject is or may readily be ascertained by the investigator. Identifiers include any and all HIPAA identifiers.

7.10. **IDENTIFIABLE BIOSPECIMEN:** A biospecimen for the identity of the subject is or may be readily ascertained by the investigator.

7.11. **LIMITED DATA SET:** A data set that excludes certain direct identifiers but may include other identifiers (e.g. dates) that require a data use agreement for disclosure.

7.12. **MATERIAL TRANSFER AGREEMENT:** A contract that governs the transfer of tangible research materials between two organizations when the recipient intends to use it for his or her own research purposes. The MTA defines the rights of the provider and the rights and obligations of the recipient with respect to the materials and any progeny, derivatives, or modifications. The material may take any form, from chemicals, electronics, biologicals, or any other material.

7.13. **NON-HUMAN SUBJECTS RESEARCH:** Activities that do not meet the definitions of human subjects and/or research as defined in the Code of Federal Regulations 45 Part 46 titled Policy for Protection of Human Research Subjects. Non-human research involves only anonymized data and/or specimens.

7.14. **PSEUDONYMISATION:** For data subject to General Data Protection Regulation (GDPR), replacing any information which could be used to identify an individual with a pseudonym, or a value which does not allow the individual to be directly identified, but for which a key may still exist at another location.

7.15. **REPOSITORY: T**he storage site or mechanism by which data/biospecimens are collected from one or more sources and stored or archived in a form designed for ease and speed of aggregation, search, and retrieval. Repositories include data banks, tissue banks, and registries that collect, store, and distribute data/biospecimens for use in current and future research projects.

7.16. **SHARING AGREEMENT:** An agreement that details the conditions for receipt and future use of data/biospecimens from a repository for when Material Transfer Agreement or Data Use Agreements are not required.

7.17. **SUBMITTAL AGREEMENT:** An agreement that attests that human data/biospecimens were obtained with written informed consent of the participant using an informed consent document approved by the local IRB, or pursuant to an IRB-approved waiver of informed consent.

8. REFERENCES

- Broad Consent: 45 CFR 46.116(d)
- Human Subject: 45 CFR 46.102 (e)(1)
- Identifiable Private Information: 45 CFR 46.102(e)(5)
- Identifiable Biospecimen: 45 CFR 46.102(e)(6)
- Limited Data Set: HIPAA Privacy Rule
- U.S. Department of Health and Human Services, Coded Private Information or Biospecimens Used in Research: Guidance (2018), https://www.hhs.gov/ohrp/regulations-and-policy/guidance/research-involving-coded-private-information/index.html.; 45 CFR 46.102(e)(1).
- U.S. Department of Health and Human Services, Office for Human Research Protections. Exempt Research Determinations. Accessed May 3, 2025. https://www.hhs.gov/ohrp/regulations-and-policy/guidance/exempt-research-determination/index.html.
- Wade TD. Traits and types of health data repositories. Health Inf Sci Syst. 2014 Jun 30;2:4. doi: 10.1186/2047-2501-2-4. PMID: 25825668; PMCID: PMC4340801; University of Southern California. Policy on Biorepositories. Accessed May 3, 2025. https://policy.usc.edu/biorepositories/.
- 45 CFR 46.104
- University of Southern California. Policy on Biorepositories. Accessed May 3, 2025. https://policy.usc.edu/biorepositories/.
- Sharing Data for Research Purposes guide, Yale HIPAA Privacy Office

9. RESOURCES

9.1. Appendix 1: Repository Submission Tracking Tool
9.2. Appendix 2: Repository Release Tracking Tool
9.3. Appendix 3: Submittal Agreement Template
9.4. Appendix 4: Sharing Agreement Template

Yale University
Human Research Protection Program (HRPP)
P.O. Box 208327
New Haven, CT 06520-8327
Phone: 203-785-4688

APPENDIX 1: REPOSITORY SUBMISSION TRACKING TOOL

Source of Data/Biospecimens	Description of Data/Biospecimens	Level of Identifiability	Consent and/or HIPAA Authorization	Submittal Agreement Execution Date	Transaction Date
Study ID #: PI Name:		☐Fully Identifiable ☐Coded ☐Deidentified/Anonymized ☐Anonymous	*[describe any restrictions]*		
Study ID #: PI Name:		☐Fully Identifiable ☐Coded ☐Deidentified/Anonymized ☐Anonymous	*[describe any restrictions]*		
Study ID #: PI Name:		☐Fully Identifiable ☐Coded ☐Deidentified/Anonymized ☐Anonymous	*[describe any restrictions]*		
Study ID #: PI Name:		☐Fully Identifiable ☐Coded ☐Deidentified/Anonymized ☐Anonymous	*[describe any restrictions]*		

Yale University
Human Research Protection Program (HRPP)
P.O. Box 208327
New Haven, CT 06520-8327
Phone: 203-785-4688

APPENDIX 2: REPOSITORY RELEASE TRACKING TOOL

Recipient Information	IRB Approval Status	Description of Data/Biospecimens	Level of Identifiability	Consent and/or HIPAA Authorization	Sharing Agreement Execution Date	Transaction Date
PI Name: Institution:	☐Approved ☐NHSR Determination ☐N/A non-identifiable		☐Fully Identifiable ☐Coded ☐Deidentified/Anonymized ☐Anonymous	***[describe any restrictions]***		
PI Name: Institution:	☐Approved ☐NHSR Determination ☐N/A non-identifiable		☐Fully Identifiable ☐Coded ☐Deidentified/Anonymized ☐Anonymous	***[describe any restrictions]***		
PI Name: Institution:	☐Approved ☐NHSR Determination ☐N/A non-identifiable		☐Fully Identifiable ☐Coded ☐Deidentified/Anonymized ☐Anonymous	***[describe any restrictions]***		
PI Name: Institution:	☐Approved ☐NHSR Determination ☐N/A non-identifiable		☐Fully Identifiable ☐Coded ☐Deidentified/Anonymized ☐Anonymous	***[describe any restrictions]***		

Yale University
Human Research Protection Program (HRPP)
P.O. Box 208327
New Haven, CT 06520-8327
Phone: 203-785-4688

APPENDIX 3: SUBMITTAL AGREEMENT

[NAME OF REPOSITORY] SUBMITTAL AGREEMENT

I, the Submitting Investigator, submit the following data/biospecimens to the Repository Investigator for storage, maintenance, and future research use:

[insert description of data/biospecimens submitted to the Repository]

This Submittal Agreement documents that I agree to the stipulations of submission, as specified below **[select only as applicable]:**

☐ Data/biospecimens were originally collected for clinical purposes (e.g., clinical discards).

☐ Data/biospecimens were collected under an IRB-approved protocol.

☐ Data/biospecimens were collected under an IRB-approved consent form.

☐ A copy of the IRB-approved consent form has been provided to the Repository Investigator.

☐ There are no restrictions on the future uses of the data/biospecimens.

☐ The following restrictions exist on the future uses of the data/biospecimens:

[insert description of restrictions]

☐ Data/biospecimens were collected under an IRB waiver of consent.

☐ I confirm that, if contacted, I will not provide any subsequent recipient of the data/biospecimens with access to the identities of participant or to information through which the identities of participants may be ascertained.

☐ Data were collected prior to April 13, 2003, and HIPAA authorization is not required.

☐ Data were collected under a HIPAA Authorization or waiver requiring identified data destruction after:

[insert data destruction date]

Yale University
Human Research Protection Program (HRPP)
P.O. Box 208327
New Haven, CT 06520-8327
Phone: 203-785-4688

☐ Research participants who refused to allow data to be used for future research or who wanted to be contacted prior to future use are:

☐ Excluded from the material provided to the Repository **OR**

☐ Flagged to indicate that prior permission for future use must be obtained.

☐ I give my assurance that the data submission is accurate to the best of my knowledge.

Submitting Investigator Printed Name	Submitting Investigator Signature	Date Signed

Submitting Investigator Institution Name	Submitting Investigator Institution FWA# (if external)

Yale University
Human Research Protection Program (HRPP)
P.O. Box 208327
New Haven, CT 06520-8327
Phone: 203-785-4688

APPENDIX 4: SAMPLE SHARING AGREEMENT

[NAME OF REPOSITORY] SHARING AGREEMENT

The Recipient acknowledges that the conditions for use of the data/biospecimens are governed by the Yale University Institutional Review Board (IRB) in accordance with applicable federal, state, and local regulations and Institutional policies and procedures.

The Recipient agrees to comply fully with all such conditions and report to **[INSERT Repository Investigator Name]** any proposed changes in the recipient's research and any unanticipated problems involving risks to subjects or others in support of prompt reporting requirements. The Recipient remains subject to applicable laws, regulations, and institutional policies that protect human subjects.

Description of data/biospecimens to be shared:

[insert description of data/biospecimens]

For sharing of **identifiable** data/biospecimens:

The data/biospecimens provided to the Recipient may only be utilized in accordance with the conditions stipulated in this Agreement, as approved by the Yale University IRB, as follows:

- The Recipient assures that IRB approval has been or will be obtained for each individual research project involving the identifiable data/biospecimens.
- The Recipient will keep all participant information confidential and comply with all applicable privacy and security laws and regulations.
- The Recipient will maintain any coding system designated by the Repository to protect participant privacy and confidentiality. The key to the code will be kept securely stored in a location separate from the coded data/biospecimens.
- The Recipient will not attempt to contact the participants who originally provided the data/biospecimens.
- Any research use of the data/biospecimens beyond the scope of this Agreement is prohibited.

Yale University
Human Research Protection Program (HRPP)
P.O. Box 208327
New Haven, CT 06520-8327
Phone: 203-785-4688

For sharing of **anonymized/deidentified** data/biospecimens:

- Repository personnel **will not** provide identifying information to the Recipient under any circumstances.
- The Recipient will not attempt to identify the participants and may not contact individuals who are collecting the data/biospecimens (e.g., research assistants) to obtain any identifying information.
- The Recipient will not merge the data with another dataset unless IRB approval is obtained for that activity.
- All data/biospecimens are labeled with a code number that is assigned by the Repository for tracking purposes. The code does not include any identifying information (e.g., birthdate).
- Participant information will be kept confidential.

The data/biospecimens **may not** be used for the following types of research activities, which are prohibited by the participant's prior consent and/or HIPAA Authorization restrictions:

[insert description of restrictions]

Repository Investigator Printed Name	Repository Investigator Signature	Date Signed
Requesting Investigator Printed Name	Requesting Investigator Signature	Date Signed