OSUCCC Retrospective Data Study Tip Sheet

The purpose of this document is to provide guidance on creating a protocol for a retrospective data study. It outlines the essential components of a study protocol, includes standard language that can be used, and addresses Institutional Review Board (IRB) considerations.

Following these recommendations can help facilitate a streamlined review process, minimize the need for amendments, and ensure that study activities are conducted in compliance with university policy and state and federal regulations.

Please note: In Spring 2025, the IRB moved to a new system, Huron, and issued new protocol templates. The language in this document can still be utilized towards the new protocol template.

I. Retrospective Data Study Protocol

Consider the following details when developing a protocol.

Sample size and approach:

Explicitly indicate the number of participants expected for the study. This could be based on an average number of patients seen in clinics or an aggregate count obtained through the OSUWMC honest broker process, a centralized repository (e.g., Total Cancer Care, Hematology Tissue Bank), Slicer Dicer (EPIC), or Deep 6 AI.

Waiver of Consent Process and Full Waiver of HIPAA Authorization:

State that a full waiver of HIPAA authorization is being sought as there is no direct patient contact with retrospective protocols. See Protected Health Information and HIPAA Policy.

Data Collection:

List the data elements and medical information to be collected (e.g., diagnosis, stage, comorbidities, labs, images, etc.). A data collection form or list can be included in this section, in an appendix, or as an attachment. The IRB will seek a definitive list of data elements, not data "such as," including," or "not limited to." Data collection should only be for the data elements needed to perform the research. See Minimum Necessary Requirement.

Indicate if data will contain protected health information (IRB required) or be de-identified or coded-limited. If PHI, specify whether the study team will de-identify patient identifiers such as name and MRN. Provide justification for the study maintaining PHI.

Describe how where data will be stored, who will have access to the data, and how the data is protected.

Multi-site: If other sites are providing information or if OSU is providing information to other sites, state that data will be either de-identified or coded prior to receipt/transmission. Provide the method of transmission (e.g., entered into a centralized REDCap database or other electronic data capture tool, sent via secure file transfer protocol (sftp), encrypted email, etc.) See Email Security and Encryption.

Risk and Benefits:

Generally, retrospective studies will not provide benefit to study participants but may provide benefit to future patients based on findings. Specify how patient privacy and clinical data will be protected.

Data Analysis Plan:

Detail the statistical analysis plans, software used (if appropriate), and whether other groups will be involved (e.g., Biostatistics Shared Resource). Indicate when data collection and analysis are expected to be completed.

II. Example Language:

Sample Size and approach:

- "(Number) patients are expected to match study criteria. This is based on an average (number) cases being seen in clinic per year and the (number)-year study period."
- "We plan to review (number) charts within the study date range. Data for the sample size request were obtained through the OSUWMC honest broker process."

Research Design

- "This project is retrospective in nature, as we plan to utilize existing clinical, pathologic, imaging, and treatment-related records. The insights obtained from this study will help generate hypotheses for future management of (condition/diagnosis)."
- "The use or disclosure of protected health information involves no more than a **minimal risk** to the privacy of individuals. We include adequate plans to protect the identifiers from improper use and disclosure, e.g., destroy the identifiers at the earliest opportunity consistent with research standards and PHI will not be reused or disclosed to others."

Waiver of Consent and HIPAA Authorization:

- "We will seek a waiver of consent for this research project as there is no direct patient contact in this retrospective study. Due to the date range/number of potentially eligible patients, many individuals may have since expired or are may no longer be seen by a provider of the James.

- "The research could not practicably be conducted without access to and use of PHI."

Data Collection:

- "Medical record number (MRN), demographic data (including age, sex, race/ ethnicity, height, weight, body mass index), comorbidities (e.g., diabetes, heart failure, hypertension, respiratory disease), cancer primary diagnosis, stage, histology (carcinoma, adenocarcinoma, etc.), cancer treatment data (e.g., chemotherapy, surgery, radiation, immunotherapy, hormone therapy), adverse event (AE) symptoms, PET or MRI imaging, lab results."
- "All data will be de-identified, and patient identifiable information will be replaced with a unique patient identifier (series of numbers/letters), which will then become how a participant is referenced during or after data analysis. The coding key to reverse identify patients will be stored separately from the dataset."
- "The necessity of maintaining PHI stems from the complexity of the relationships between the clinical data and the trajectory of treatment. These files will be stored on a secured hard drive where only the PI and investigators directly involved in the project will have access to the raw data."
- Multi-site: "All data (received from/sent to) external institutions will be provided without identifying information."
- Indicate study duration: "Data collection should be completed within (number) months of approval by the IRB. Analysis and reporting of the data should not take longer than (number) months beyond that point."

Risks and Benefits:

- "This study will not add any further benefit to the study participants as the data analysis consists of existing clinical health information and is retrospective in nature. However, the study has the potential to help future patients through the identification of factors and aides for the prediction, early detection, and management of (condition/diagnosis)."
- "The data will be maintained in a dedicated folder within the department shared drive that will be password protected and behind the OSUWMC firewall. Only authorized study personnel reflected on the protocol will have access to the data."
- "After the data collection process is complete the data will be de-identified."
- "Data will be destroyed five years (university policy) after study completion." See Records Management Policy.
- "No personal identifiers will be used to reference participants during the study, and a unique patient identification number will be used instead."
- "The confidentiality of each patient record will be rigorously maintained using existing standards at The Ohio State University Health Systems. Health Insurance Portability and Accountability Act (HIPAA) and state/federal government regulations for protecting patient privacy and security will

be strictly maintained. The results of the research studies may be published but subjects will not be identified in any publication." See OSU-IT Security Policy.

- "When personnel leave the project or unit, user access rights will be terminated."
- "All those listed in the protocol who can access the data have completed HIPAA and CITI Human Subjects Research training."

III. Retrospective Data Study Tips

Note: *Two or more active retrospective protocols cannot be combined to conduct research activities.*

Multi-site study:

- If multi-site, list all currently known participating **research locations** and indicate if Ohio State or another institution will be the lead site, Institutional Review Board (IRB) of record, and/or data management and coordinating center.

Cohort Identification:

- The **study cohort** should be well-defined.
- Note that medical records accessed for patients found to be ineligible *still count* toward the total number of approved participants, even if cases are not used in the final analysis.
- Detail the procedures used to **identify** and **screen** participants who meet eligibility criteria.
    - Are participants identified through an honest broker process, manual chart review, physician referral, departmental databases, or a combination of these sources?
    - An aggregate number of participants matching study inclusion criteria can be provided through the OSUWMC or centralized CCC repository honest broker processes (e.g., Total Cancer Care, Hematology Tissues Bank), Slicer Dicer, or Deep 6 AI to **assess feasibility** for grants, protocol enrollment projections, and justification for participant numbers.
    - If participants are identified from an existing **research database**, include the OSU study or IRB number for that database protocol.
    - **Vulnerable patient populations** that may need to be excluded from research include minors, patients with impaired decision-making abilities, and prisoners.
    - If **patient controls** are used for comparison, what are the criteria for this group, and will the same or additional data elements be collected from them?

Waiver of Consent Process and Full Waiver of HIPAA Research Authorization:

- A **waiver of consent** is the most frequently requested and utilized approach in a retrospective study conducting chart review. For the IRB to waive the consent process, the following criteria must be met:
    - research involves no more than **minimal risk**,
    - the waiver or alteration will not adversely affect participant **rights** and **welfare**, and
    - research could not practicably be carried out without the waiver or alteration.

- Retrospective protocols require a **full waiver** of HIPAA research authorization to view and collect existing patient health information from within the electronic medical record (IHIS).

Data Collection:

- Provide detailed data **collection methods** (including how to avoid and minimize participant risks) so that the IRB can assess the potential study risks and benefits. See Research Data Policy.
- Will data be obtained through manual review of the electronic medical record (IHIS) by the study team or electronic extraction by an honest broker process, or both?
    - Regardless, data provided must be within the IRB-approved **date range** reflected in the protocol and the full waiver of HIPAA authorization appendix in Huron.
    - If seeking outcomes data, consider extending the study date range to include this period.
- PHI from **non-Ohio State locations** and hospital systems cannot be included in analysis unless necessary approvals are in place for the institution to serve as a participating research site.
    - This includes Care Everywhere, OSUWMC affiliates, and Community Connect sites (e.g., Adena Health, Community Memorial Hospital, Mercer Health, etc.).
- Describe the methods that will be used to maximize the completeness and accuracy of data.
- Individual **PHI elements** and **dates** (DOB, diagnosis, procedures, etc.) should be listed in the protocol and under the full waiver of HIPAA authorization section.
- Will data eventually be **coded** or **de-identified**? Include the timepoint during the study at which this would occur.
- Consider whether **images** (not just reports) are needed, and if so, will they require de-identification. This process requires research imaging informatics services and may carry a fee.
- Where will data/PHI be **stored**? Include the building, room number(s), and/or specific server.
    - Examples: password-protected Excel spreadsheet, REDCap or another platform/software, physical papers in a locked cabinet within the PI's office, shared folder or department drive behind the medical center firewall, etc.
- Indicate how long participant information will be stored once the study is completed.
    - University policy states that data must be retained for a **minimum of five years**. See Records Management Policy.

Data Transfers:

- Data entering or leaving Ohio State, which is typically a coded-limited or de-identified dataset, requires a reviewed and signed **Data Use Agreement** (DUA) to be in place with the Office of Innovation and Economic Development (formerly "TCO") or the Office of Legal Affairs before information can be shared.
- University data must be transferred in a **secure manner,** including an electronic data capture system like REDCap, encrypted email, secure file transfer protocol (sftp), etc.

Risks and Benefits:

- *All* retrospective studies are associated with a risk to **breach of data confidentiality**. What practices or processes are in place to minimize the loss of privacy?

- Examples include user access granted (password), only approved personnel will have access to study information, replacing MRNs with unique study IDs (random or sequential numbers/letters), storing electronic records on a shared drive, etc.

- The loss of confidential research data must be **reported** to the IRB as it poses a potential risk to study participants. Potential breaches of confidentiality involving PHI must also be reported to the applicable HIPAA privacy officer as soon as possible.

- Retrospective study subjects do not receive direct **patient benefits**, however, what is learned may increase general knowledge and influence the treatment of future patients.

Ohio State Contact Information

Office of Responsible Research Practices (ORRP, IRB staff):

- https://research.osu.edu/contact-us?department=4
- General phone: (614)688-8457

OSU Information Technology (OSU-IT):

- https://it.osu.edu/help
- Medical Center: (614)292-6446 (HELP)
- University: (614)688-6446 (HELP)

Office of University Compliance and Integrity (OUCI):

- https://compliance.osu.edu/
- General phone: (614)292-3251

Ohio State Policy and Guidance

Institutional Data Policy:

- https://policies.osu.edu/sites/default/files/documents/2025/02/institutional-data-policy.pdf

Protected Health Information and HIPAA (OSUWMC):

- https://policies.osu.edu/sites/default/files/documents/2025/04/Protected-Health-Information-HIPAA.pdf

Research Data

- https://policies.osu.edu/sites/default/files/documents/2025/03/research-data-policy.pdf

Research Health Information ("RHI," OSU-OTDI):

- https://it.osu.edu/security/research-support/research-health-information

Research Recruitment Tip Sheet for Investigators (OSU-COM):

- https://medicine.osu.edu/-/media/files/medicine/research/office-of-research/20230217-gdl-inst-113-res-recruitment-tip-sheet.pdf?rev=cbd55b39da20408b8e1e55748bad5973

Research Using Protected Health Information ("PHI," OSU-ERIK):

- https://research.osu.edu/research-responsibilities-and-compliance/human-subjects/hipaa-and-human-subjects-research/research

Secure Email (OSUSecure):

- https://it.osu.edu/security/services/secure-email

Ohio State Resources

OSUCCC Hematology Tissue Bank (HTB) Requests:

- HTB request form: https://go.osu.edu/htbrequest

OSUCCC Total Cancer Care (TCC) Data Requests:

- TCC request form: https://go.osu.edu/osutcc

OSUWMC Honest Broker Research Data Requests:

- Research Data Request form: https://go.osu.edu/researchdatarequest

US Federal Guidance

Definition for Coded-Limited Data Set:

- https://www.hhs.gov/hipaa/for-professionals/special-topics/emergency-preparedness/limited-data-set/index.html#:~:text=A%20LDS%20is%20protected% 20health,Telephone%20numbers

Health Information Privacy:

- https://www.hhs.gov/hipaa/for-professionals/special-topics/research/index.html

Methods for De-Identification of PHI:

- https://www.hhs.gov/hipaa/for-professionals/special-topics/de-identification/index.html

Minimum Necessary Requirement:

- https://www.hhs.gov/hipaa/for-professionals/privacy/guidance/minimum-necessary-requirement/index.html

Protecting Personal Health Information in Research: Understanding the HIPAA Privacy Rule

- https://privacyruleandresearch.nih.gov/pdf/HIPAA_Privacy_Rule_Booklet.pdf