TEXAS A&M UNIVERSITY
SAN ANTONIO

Institutional Review Board

GUIDELINE #5: RESEARCH INVOLVING SECONDARY USE OF DATA, DOCUMENTS, RECORDS OR SPECIMENS

I. PURPOSE

This guideline is to ensure that human subjects' research conducted complies with federal, state, and local laws, regulations, directives, and instructions.

This guideline applies only to protocols that involve the secondary analysis of existing data, such as medical records, student records, data collected from previous studies, audio/video recordings, etc. that were initially collected for another purpose. Though such protocols do not involve interactions or interventions with humans, they may still require Texas A&M University – San Antonio Institutional Review Board (A&M-SA IRB) review, since the definition of "human subject" at 45 CFR 46.102(f) includes living individuals about whom an investigator obtains identifiable private information for research purposes.

Data analysis activities that meet the definition of research with human subjects may qualify for an exemption or require expedited or even full committee review. Any such protocol must receive A&M-SA IRB approval or a determination of exemption before the investigator accesses the data.

II. STATEMENT

All human subjects research, irrespective of the source of funding, conducted by A&M-SA faculty, staff and students must be submitted and reviewed in accordance with Federal research regulations, Texas A&M System Guidelines, A&M-SA IRB policies and local consideration.

III. SCOPE

This guideline applies to all research conducted where the A&M-SA IRB serves as the Reviewing IRB.

IV. PROCEDURE

When does the secondary use of existing data not require review?

In general, the secondary analysis of existing data does not require A&M-SA IRB review when it does not fall within the regulatory definition of research involving human subjects, as referenced above.

Note: Although the definition of a human subject includes only living individuals, thereby excluding decedents, there are cases in which the health information of the deceased and death data files may require A&M-SA IRB review.

Public data: Public use data sets (such as portions of U.S. Census data, data from the National Center for Educational Statistics, National Center for Health Statistics, etc.) are data sets prepared with the intent of making them available for the public. The data available to the public are not individually identifiable and

therefore their analysis would not involve human subjects.

In addition to being identifiable, the existing data must include "private information" to constitute research involving human subjects. Private information is defined as information which has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public (e.g., a medical or school record). Information that contains identifiers and can be accessed freely by the public (without special permission or application) is not "private" and the research therefore does not therefore involve human subjects. For example, a study involving only analysis of the published salaries and benefits of public university presidents would not need IRB review since this information is not private.

De-identified data:

If the dataset has been stripped of all identifying information and there is no way to be linked back to the subjects from whom it was originally collected (through a key to a coding system or by any other means), its subsequent use by the PI or another investigator would not constitute human subjects research, since it is no longer identifiable.

Identifiable means the identity of the subject is known or may be readily ascertained by the investigator or associated with the information. In general, information is identifiable when it can be linked to specific individuals by the investigator(s) either directly or indirectly through coding systems, or when characteristics of the information obtained are such that by their nature a reasonably knowledgeable person could ascertain the identities of individuals. Therefore, even though a dataset may have been stripped of direct identifiers (names, addresses, student ID numbers, etc.), it may still be possible to identify an individual through a combination of other characteristics (e.g., age, gender, ethnicity, and place of employment).

A&M-SA IRB will confirm that the research does not require IRB review.

Scenario:

1. Many student research projects involve secondary analysis of data that belongs to, or was initially collected by, their faculty advisor or another investigator. If the student is provided with a de-identified, non-coded data set, the use of the data does that constitute research with human subjects?

Key: No. No individuals are being contacted; no private information is obtained.

Coded data:

Secondary analysis of coded private information is not considered to be research involving human subjects and would not require A&M-SA IRB review if the investigator(s) cannot readily ascertain the identity of the individual(s) to whom the coded private information pertains as a result of one of the following circumstances:

- The investigators and the holder of the key have entered into an agreement prohibiting the release of the key to the investigators under any circumstances, until the individuals are deceased (DHHS regulations for human subjects research do not require the IRB to review and approve this agreement);

- There are IRB-approved written policies and operating procedures for a repository or data management center that prohibit the release of the key to the investigator under any circumstances, until the individuals are deceased; or
- There are other legal requirements prohibiting the release of the key to the investigators, until the individuals are deceased.

Considerations:

- If a student is analyzing coded data from a faculty advisor/sponsor who retains a key, this would be human subjects research, because the faculty advisor is considered an investigator on the student's protocol and can readily ascertain the identity of the subjects since he/she holds the key to the coded data.
- If the student's work fits within the scope of the initial protocol from which the dataset originates, the faculty advisor (or investigator who holds the dataset) may wish to consider adding the student and his/her work to the original protocol by means of an amendment application rather than having the student submit a new application for review.

Scenario

1. Researcher A plans to examine the relationships between attention deficit hyperactivity disorder (ADHD), oppositional defiance disorder, and teen drug abuse using data collected by Agencies A, B, and C that work with "at risk" youth. The data will be coded, and the agencies have entered into an agreement prohibiting release of the key to the researcher that could connect the data with identifiers. Would the use of the data not constitute research with human subjects and does not require IRB review?

Key: Not Human Subjects Research (NHSR), A&M-SA requires NHSR certification for these studies.

When is the secondary use of existing data exempt?

The six categories of research activities involving human subjects that may be exempt from the requirements of the Federal Policy for the Protection of Human Subjects (45 CFR 46) are presented in exemption categories 1 to 6.

However, only one category (Category 4) applies specifically to secondary data analysis. If research is found to be exempt, it need not receive full or subcommittee (expedited) review. To qualify for an exempt determination, a Protocol application must be submitted to A&M-SA for review.

Research involving secondary data analysis of data, documents, and biospecimens can be exempted under Category 4 of the federal regulations if:

- the sources of such data are publicly available; or
- the information is recorded by the investigator in such a manner that the resulting dataset contains no information that can identify subjects, directly or through identifiers linked to the subjects.

The latter condition of this category applies in cases where the investigators initially have access to identifiable private information but abstract the data needed for the research in such a way that the information can no longer be connected to the identity of the subjects.

This means that the abstracted data set does not include direct identifiers (names, social security numbers, addresses, phone numbers, etc.) or indirect identifiers (codes or pseudonyms that are linked to the subject's identity). Furthermore, it must not be possible to identify subjects by combining a number of characteristics (e.g., date of birth, gender, position, and place of employment). This is especially relevant in smaller datasets, where the population is confined to a limited subject pool. Data must be de-identified before any analysis is conducted to qualify under exempt category 4.

The following do not qualify for exemption:

- Research involving prisoners except for research aimed at involving a broader subject population that only incidentally includes prisoners,
- research involving protected health information from HIPAA-covered entities, and
- FDA-regulated research.

Scenarios:

1. A researcher conducts a study of treatment outcomes for a certain drug that involves the review of patient charts at a non-A&M-SA facility. The researcher records patient age, sex, diagnosis, and treatment outcome in such a way that the information cannot be linked back to the patient. How should this protocol be reviewed?

Key: This project could qualify for an exemption.

2. Student B will be given access to data from her faculty advisor's health survey research project. The data consists of coded survey responses, and the advisor will retain a key that would link the data to identifiers. The student will extract the information she needs for her project without including any identifying information and without retaining the code. What level of review is required for this study?

Key: The use of the data does constitute research with human subjects because the initial data set is identifiable (albeit through a coding system); however, it would qualify for exempt status.

When is the secondary use of existing data non-exempt?

If secondary analysis of existing data does involve research with human subjects and does not qualify for exempt status as explained above, the project must be reviewed either through expedited procedures or by a full (convened) Board, and a non-exempt application must be submitted for A&M-SA IRB review.

Consent waivers:

Researchers using data previously collected under another study should consider whether the currently proposed research is a "compatible use" with what subjects agreed to in the original consent form. For non-exempt projects, a consent process description or justification for a waiver must be included in the research protocol.

IRB may require that informed consent for secondary analysis be obtained from subjects whose data will be accessed. Alternatively, A&M-SA IRB can consider a request for a waiver of one or more elements of informed consent under 45 CFR 46.116(f).

- In order to approve such waiver, the IRB must first be satisfied that:
  the research presents no more than minimal risk of harm to the subjects; and
- the research could not practicably be carried out without the waiver or alteration; and
- if the research involves using identifiable private information or identifiable biospecimens, the research could not practicably be carried out without using such information or biospecimens in an identifiable format;*
- the waiver or alteration will not adversely affect the rights and welfare of the subjects; and whenever appropriate, the subjects or legally authorized representatives will be provided with additional pertinent information after participation.

  * If an individual was asked to provide broad consent for the storage, maintenance, and secondary research use of identifiable private information or identifiable biospecimens, and refused to consent, an IRB cannot waive consent for the storage, maintenance, or secondary research use of the identifiable private information or identifiable biospecimens.

"Restricted Use Data":

- Certain agencies and research organizations release files to researchers with specific restrictions regarding their use and storage.
- The records frequently contain identifiers or extensive variables that combined might enable identification, even though this is not the intent of the researcher.
- Research using these data sets most often requires non-exempt level review.

Scenarios:

1. Student C will be given access to coded mental health assessments from his faculty advisor's research project. The student plans to analyze the data with a code attached to each record, and the advisor will retain a key to the code that would link the data to identifiers. What level of review is required by IRB?

Key: While the use of the data does constitute research with human subjects, it does not qualify for exempt status since subjects can be identified. This student project would require an application to be submitted for non-exempt review by the IRB

2. The PI of the aforementioned study requests IRB for ideas to reduce regulatory burden that will allow the student to complete the dissertation on time. What should IRB do?

Key: If the student's work fits within the scope of the initial protocol from which the dataset originates, the faculty advisor (or investigator who holds the dataset) may consider adding the student and his/her work to the original protocol by means of an amendment application rather than having the student submit a new application for non-exempt review.

3. Student D is applying to the National Center for Health Statistics for use of data from the National Health and Nutrition Examination Survey that includes geographic identifiers and date of examination. Should IRB review the study?

Key: The analysis of this restricted use data would require non-exempt review by IRB.

When is secondary data (e.g., medical records, purchased data, data from the Internet, etc.) considered human subjects research?

Research involving secondary data analysis is considered human subjects research when data about individuals is both private and identifiable. Researchers who are unsure whether their project fits under this category should contact A&M-SA IRB for consultation.

Projects that might be human subjects research because they involve:

• Purchasing/obtaining enhanced data sets—data on individuals which may include enough information to potentially identify the individuals.
• Receipt of coded data where data holder has code key—depending on whether the data holder only provides data or is a collaborator in the research, and whether an agreement between institutions prohibits receiver from ever receiving identifiers, etc.
• Forums or chats where users must register as belonging to a certain group (e.g., cancer survivors) or housed in areas that are not public, e.g., where special passwords are needed to join.

Projects that are human subjects research because they involve:

• Private data sets obtained with identifiers (e.g., traffic violation data with driver's license numbers, survey data with email addresses, medical records with protected health information [PHI], restricted use datasets, etc.).
• Stolen, hacked, accidentally released data about individuals—although data may now be publicly available (such as on the surface web or the dark web), the individuals whom the data is about had expectation of privacy, i.e., that the data would not be hacked, stolen, etc.

Human subjects research must be reviewed and either determined exempt or obtain IRB approval before the research can begin.

Examples of projects that are unlikely to be human subjects research because they involve only:

• Public use data sets such as data from the National Center for Health Statistics—data is available to the public at large and not restricted to researchers.
• Data sets from an outside source that have been stripped of all identifying information and of links back to identifiers before being provided to researcher.
• Facebook public profiles found from Google searches.

• X, formally known as Twitter tweets not in private setting.
• Publicly accessible forums or comments sections where users have no expectation of privacy (e.g., New York Times, YouTube, etc.).

V. APPLICABLE REGULATIONS AND GUIDELINES

Office of Human Research Protections: Coded Private Information or Specimens Use in Research, Guidance (2008)

VI. REVISIONS

March 2023
October 2023

Guideline: # 5	**Version:** V.2
Title: Research Involving Secondary Use of Data, Documents, Records or Specimens	
Authorized: Dr. Vijay Golla, Vice Provost for Research and Health Sciences	
Date Approved by IO:	

Institutional Review Board

GUIDELINE #5 FORM: RESEARCH INVOLVING SECONDARY USE OF DATA, DOCUMENTS, RECORDS OR SPECIMENS

- *If your proposed project involves any activities other than the secondary use of data, documents, records or specimens do not use this form.*
- *Secondary research refers to the research use of information or biospecimens that were collected for another purpose such as clinical care, education records, or a different research project.*

I. PURPOSE

This guideline is to ensure that human subjects' research conducted complies with federal, state and local laws, regulations, directives and instructions.

II. STATEMENT

All human subjects research, irrespective of the source of funding, conducted by A&M- SA faculty, staff and students must be submitted and reviewed in accordance with the Federal research regulations, Texas A&M System Guidelines, A&M-SA IRB policies and local consideration.

III. SCOPE

This guideline applies to all research conducted where the A&M- SA IRB serves as the Reviewing IRB

1. Protocol Title: **Click here to enter text.**

2. Protocol Version Date: **Click here to enter text.**

3. Principal Investigator: **Click here to enter text.**

4. Other Investigators: **Click here to enter text.**

5. Funding:

Will this project be sponsored or funded:

Yes ☐ No ☐

If yes, indicate the sponsor and upload a copy of the contract or grant and include Maestro number if applicable:

Click here to enter text.

6. Background

Provide the scientific or scholarly background, rationale, and significance of the research based on existing literature and how will it add to existing knowledge.

Click here to enter text.

7. Objectives

State the hypotheses to be tested.

Describe the purpose and primary objectives of this project.

Click here to enter text.

8. Inclusion and Exclusion Criteria for Selection of Subject Data or Specimens

Inclusion Criteria:

Click here to enter text.

Exclusion Criteria:

Click here to enter text.

Age Range: **Click here to enter text.**

Provide the age range of all subjects that meet the inclusion criteria.

Existing: Does all the data, documents, records or specimens already exist at the time this study is submitted for initial IRB review?

Yes ☐ No ☐

Date Range: Provide the date range of the data, documents, records or specimens to be analyzed or collected:

From: **Month/Day/Year** **to** **Month/Day/Year**

9. Number of Records or Specimens to be requested:

Enter the number of charts/records/specimens to be reviewed or analyzed: Click here to enter text.

Provide a rationale (e.g. statistical justification) for the number requested: Click here to enter text.

10. Accessing Data, Documents, Records or Specimens

List the source(s) of all the information, data or specimens, for this project. Include all medical record or imaging systems, lab results, billing records, student records, log-books, repositories, registries, specimen, databanks or other sources. Provide links to any applicable websites.

Click here to enter text.

***Authorization of Access**

Investigators are required to comply with internal or external regulations, policies or agreements for accessing data, documents, records or specimens for research purposes. Before your project is approved by the IRB you must submit the appropriate authorization or if necessary, the official approval or any data use agreement to access the information described in this protocol. Provide information as necessary regarding your access: Click here to enter text.

11. Check any identifiable information you will be accessing, recording or disclosing:

	Identifiers	Accessing	Recording	Disclosing
1	Names or Initials.	☐	☐	☐
2	Street address	☐	☐	☐
	Town or City	☐	☐	☐
	Parish or County	☐	☐	☐
	Complete Zip Code	☐	☐	☐
3	All elements of dates (except year) related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year).	☐	☐	☐
4	Telephone numbers.	☐	☐	☐
5	Facsimile numbers.	☐	☐	☐
6	Electronic mail (email) addresses.	☐	☐	☐
7	Social security numbers.	☐	☐	☐
8	Medical record numbers.	☐	☐	☐
9	Health plan beneficiary numbers.	☐	☐	☐
10	Account numbers.	☐	☐	☐
11	Certificate/license numbers.	☐	☐	☐
12	Vehicle identifiers and serial numbers, including license plate numbers.	☐	☐	☐
13	Device identifiers and serial numbers.	☐	☐	☐
14	Web universal resource locators (URLs).	☐	☐	☐
15	Internet protocol (IP) address numbers.	☐	☐	☐
16	Biometric identifiers, including fingerprints and voiceprints.	☐	☐	☐
17	Full-face photographic images and any comparable images.	☐	☐	☐
18	Other unique identifying number, student ID number or code	☐	☐	☐

12. Data Fields

List all other data fields that will be collected from records or other sources or attach a data dictionary (Excel file) to the IRB application.

Click here to enter text.

13. Protection of Data, Documents, Records or Specimens

Indicate how the data, documents, records or specimens that you obtain and/or the study information you record will be adequately protected from improper use and disclosure.

☐ All electronic data and recorded information will be on an authorized system computer requiring a password for access.

☐ Private personal computers, laptops and portable devices will not be used to store or access the data.

☐ All electronic data and recorded information will be encrypted.

☐ All paper records will be stored in a locked room/file-cabinet with access limited to the study team to those individuals listed as investigators on this form.

☐ All specimens will be stored in a secured/locked lab or freezer with access limited to the members indicated in this application.

☐ I will obtain a Certificate of Confidentiality; explain: Click here to enter text.

☐ Other - Describe Click here to enter text.

☐ **By checking this box, the principal investigator is acknowledging that study records will be maintained in accordance with Texas A&M System Regulations and Texas A&M University-San Antonio Rules. HIPAA authorizations signed and dated by subjects must be maintained for at least six years after completion of the research.*

14. Explain how data or specimens will be transported or transmitted:

Click here to enter text.

15. Destruction of Personal Identifiers

Indicate the earliest opportunity you will use to destroy all personal identifiers obtained and recorded:

☐ Upon data entry and validation

☐ At completion of data analysis

☐ At completion of specimen processing

☐ If there are no plans to destroy the identifiers please provide justification below;

Please explain: Click here to enter text.

16. Will there be a link, code or any other process that will allow you to connect your study data back to the identity of the subject? Yes ☐ No ☐

If yes, identify the person(s) that will maintain the link or code: Click here to enter text.

17. Data or Specimen Banking for Future Use
Data or specimen banking for future use is not to be confused with holding data or specimens for analysis at a later time for this project.

If congruent with any applicable consent, authorization or agreement, will identifiable data or identifiable specimens be banked for future use in another research project?
☐ Yes ☐ No

If yes, describe where the data or specimens will be stored, how any personal identifiers will be maintained and the procedures to release data, including: the process to request a release, approvals required for release, who can obtain data or specimens, and the data to be provided with specimens. (A separate repository protocol may be required).
Click here to enter text.

18. Other Procedures Involved
Describe any additional procedures and methods involved in your study as applicable:
Click here to enter text.

19. Study Timelines
Provide the estimated date for the investigators to complete this study:
Click here to enter text.

20. Risks to Subjects
Research involving data, documents, records or specimens pose the risk of loss of confidentiality. Are there any other risks besides breach of confidentiality?
Yes ☐ No ☐
If 'Yes' please describe: Click here to enter text.

21. Potential Benefits
Describe the potential benefits to subjects, science and/or society that may accrue as a result of this research: Click here to enter text.

22. Vulnerable Populations
This research will include information on the following populations: Check all that apply:

☐ Children
☐ Pregnant Women
☐ Neonates
☐ *Prisoners
☐ Students
☐ Employees
☐ Cognitively Impaired
☐ Economically disadvantaged

23. Sharing Study Results
☐ Results will not be shared with subjects.
☐ Results will be shared with subjects as described:
Click here to enter text.

☐ Results of this study will be published.
☐ Results of this study will be presented at the following:
Click here to enter text.
☐ Results of this study will be shared with the following drug or device companies or other industry-related sponsors (attach relevant documentation). Click here to enter text.

24. Locations and Sites
List all sites or locations where this research will be conducted. Indicate if there will be any external (non- A&M-SA) collaborators.
Click here to enter text.

Will someone other than the Principal Investigator listed on this protocol be responsible for the research listed at the sites above?
Yes ☐ No ☐
If 'Yes', please provide details: Click here to enter text.

25. Resources Available
Describe your facilities or setting: Click here to enter text.
Describe your process to ensure that all persons assisting with the research are adequately informed about the protocol, the research procedures, and their duties and functions.
Click here to enter text.

26. HIPAA Authorization for Research
Do you want to request a waiver of authorization? ☐Yes ☐ No
If 'Yes', upload a copy of the Waiver of Authorization signed by the principal investigator.

Note: Per A&M system, you may only use the data for which individuals have provided consent. Refer to the *Guidelines for Tissue Collection and Use for more* information regarding human tissues.

Guideline: # 5	**Version: V.1**
Title: Research Involving Secondary Use of Data, Documents, Records or Specimens	
Authorized: Dr. Vijay Golla, Vice Provost for Research and Graduate Studies	
Date Approved by IO: March 29, 2023	