

The use of existing data with information on living individuals may constitute “involvement of human subjects.” Projects using such data are therefore subject to all policies designed to protect the rights, safety and welfare of research participants. In accordance with OHRP guidelines, the IRB retains final authority to determine the extent of its oversight regarding the use of existing data.

1.0 Sources of existing data or information:

Existing data refers to data or information that has already been collected for purposes other than the currently proposed research. Examples of existing data include:

- Public use data repositories
- Privately held data, information, or records
- Public records (i.e. police records, phone directories, etc.)
- Published information (i.e., newspapers, public web sites, journals, books)
- Confidential records (i.e., academic or medical records; see Section 11.1)

If the data contains information about living individuals (such as their behavior, health, activities, knowledge, opinions, etc.), its use in research may constitute “involvement of human subjects.”

2.0 Involvement of human subjects:

The involvement of human subjects in research must be considered in terms of whether the data could potentially identify living individuals, either directly or through indirect means. *Since the IRB has final authority to make a determination regarding the involvement of human subjects, a protocol is required for most projects using such data.* A protocol form as well as an attachment describing the use of additional materials should be completed.

It should be emphasized that research which uses data on deceased individuals or data already in the public domain (such as information published in a journal, newspaper, or public website) does *not* have to file a protocol with the IRB. See the definition of ‘public information’ for further discussion. However, a protocol would be required if this public information is to be combined with private information.

2.1 No human subjects involved. Obtaining coded or identified existing information for research use **would not** constitute involvement of human subjects if:

- the research *only* involves the coded or identified data or information, AND
- the investigator(s) cannot readily ascertain the identity of the individual(s) to whom the data pertains because, for example:
 - the key to decipher the code is destroyed before the research begins;
 - the investigators and the holder of the key enter into an agreement prohibiting the release of the key to the investigators under any circumstances, until the individuals are deceased;

- there are IRB-approved written policies and operating procedures for a repository or data management center that prohibits the release of the key to the investigators 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.

In addition, the use of anonymous or anonymized (de-identified) pre-existing data or information **would not** constitute involvement of human subjects if the research is limited to this information.

2.2 Human subjects involved. Obtaining coded or identified existing information for research use **would** constitute involvement of human subjects if:

- The research will also involve medical records, health information, or other private information associated with the individual(s), or
- The investigators will have access to the code or key; the investigators will attempt to contact individuals; or the investigators unexpectedly learn the identity of one or more individuals

In addition, the use of anonymous or anonymized (de-identified) pre-existing data or information **would** constitute involvement of human subjects if associated medical, health, or other information will also be collected, or the individuals will be contacted.

In each set of circumstances, the research may qualify for exemption (category #4) or expedited review (category #5), or it may require full board review. See the protocol forms for more information regarding review criteria. Informed consent must be obtained from participants, unless the IRB approves a waiver of this requirement.

3.0 Data banks or repositories

The collection of data or information about individuals with the intention of creating a data bank or repository for future research purposes would constitute the involvement of human subjects. Such a project would therefore require IRB review and approval. The research may qualify for expedited review or require a review by the convened IRB. Informed consent from the individuals would also be required, unless the criteria for a waiver can be met. Research that involves the utilization of information from a data bank or repository may also constitute involvement of human subjects. A protocol must be submitted to the IRB for a determination.

4.0 Potential risks and harms to research subjects:

Use of existing data or information for research may pose significant privacy and confidentiality concerns. If informed consent will not be sought from the affected individuals, the IRB must consider whether this violates the rights of those individuals or is otherwise inconsistent with any previous agreement or terms under which the original information was gathered. There may be situations where certain policies, terms or agreements would preclude use of the material for research purposes, even if de-identified.

Harm to groups or communities may also result, even where data or information is individually de-identified, but the group or community is named in the research results. Refer to *8.2 Privacy and Confidentiality* for more information.

Private, identifiable information may place individuals at risk in the event of an unintended breach of confidentiality. Such data requires robust safeguarding procedures. Potential risks may be related to civil or criminal liability, damage to financial standing or reputation, or employability. A breach of confidential data is considered to be an unanticipated event involving risks to subjects, and requires a report to the IRB. It may also be necessary to notify participants, so they can take appropriate measures to protect themselves. Consult the ITS Security Officer for appropriate current data security procedures for access, transfer, use and storage of electronic data. Refer to *8.2 Privacy and Confidentiality* for more information.

DEFINITIONS:

Anonymized (de-identified): identifiers were originally collected, but have been irreversibly removed from previously identified samples; individual can no longer be identified or linked with their information.

Anonymous: no identifiable information exists; individual identity cannot be known or deduced, no possibility of linkage with additional information or future data collection.

Coded: 1) identifiable information has been replaced with a number, letter, symbol, or combination thereof (i.e., the code); and 2) a key to decipher the code exists, enabling linkage of the identifying information to the private information or specimens.

Human Subject: (HHS) a living individual, about whom, an investigator (whether professional or student) conducting research obtains:

- Data through intervention or interaction with the individual, or
- Identifiable private information.

Human Subject: (FDA) an individual who is or becomes a participant in research, either as a recipient of the test article or as a control. A subject may be either a healthy human or a patient. This would include individuals whose private information is used to test the safety or efficacy of a diagnostic device, even if the information is not individually identifiable, and was obtained in a retrospective fashion.

Individually identifiable: the identity of the subject is or may be readily ascertained or associated with the information; data can be linked to specific individuals either directly or indirectly through coding systems. This would also include some demographic information, or other unique information or key details that would allow individual identification to be deduced (i.e., using internet search engines or other means).

Investigator: anyone involved in conducting the research; i.e., study design or supervision, data collection, obtaining informed consent, performing research procedures, obtaining coded private information or specimens, analyzing data (note that this would *not* include someone whose sole role is providing coded private information or specimens to an investigator)

Obtaining: receiving or accessing identifiable private information for research purposes; includes an investigator's use, study, or analysis for research purposes of identifiable private information already in the possession of the investigator

Private information: information that has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public (for example, medical records). Private information must be individually identifiable (i.e., the identity of the subject is or may readily be ascertained by the investigator or associated with the information) in order to constitute 'research involving human subjects.' *The IRB has the sole authority to determine whether or not a research project constitutes the involvement of human subjects.*

Public information: information that has been collected by a government agency (such as the federal Bureau of the Census or a state Office of the Secretary of State) or a private organization (such as a newspaper or a journal) and made available for consumption by the general public. In accessing such information, researchers have had no interactions or interventions with living individual subjects. While public information may identify living individual subjects, it does not constitute 'research involving human subjects.'

Test article: any drug (including a biological product for human use), medical device for human use, human food additive, color additive, electronic product, or any other article subject to the Federal Food, Drug and Cosmetic Act.

REFERENCES:

[45CFR46.102\(f\)](#) Definition of human subject

[45CFR46.111\(a\)\(7\)](#) Criteria for IRB approval – privacy and confidentiality protections

[45CFR46.116](#) General requirements for informed consent

[OHRP Informed Consent FAQs](#)

[OHRP Guidance on Research Involving Coded Private Information or Biological Specimens](#)

[OHRP Guidance on Data or Tissue Repositories](#)

[Expedited Review Categories](#)

[Exemption Categories \(45CFR46.101\)](#)

RELATED FORMS:

IRB Protocol Form

Exemption Protocol Form

Expedited Categories Attachment

Additional Materials Attachment

Informed Consent Waiver or Alteration Request Form

RELATED HRPP SECTIONS:

2 Applicability

7 IRB Review Process

8.2 Privacy and Confidentiality

9.3 Waiver or Alteration of Informed Consent Requirements

11.1 Use of Confidential Records