OHRP - Guidance on Research Involving Coded Private Information or Biological Specimens

NOTE: THIS GUIDANCE REPLACES OHRP'S AUGUST 10, 2004 GUIDANCE ENTITLED "GUIDANCE ON RESEARCH INVOLVING CODED PRIVATE INFORMATION OR BIOLOGICAL SPECIMENS." CLICK HERE FOR THE AUGUST 10, 2004 GUIDANCE. THIS GUIDANCE HAS BEEN UPDATED TO BE CONSISTENT WITH THE CONTENT OF OHRP'S OCTOBER 16, 2008 "GUIDANCE ON ENGAGEMENT OF INSTITUTIONS IN HUMAN SUBJECTS RESEARCH."

This guidance represents OHRP's current thinking on this topic and should be viewed as recommendations unless specific regulatory requirements are cited. The use of the word *must* in OHRP guidance means that something is required under HHS regulations at 45 CFR part 46. The use of the word *should* in OHRP guidance means that something is recommended or suggested, but not required. An institution may use an alternative approach if the approach satisfies the requirements of the HHS regulations at 45 CFR part 46. OHRP is available to discuss alternative approaches at 240-453-6900 or 866-447-4777.

Date: October 16, 2008

Scope: This document applies to research involving coded private information or human biological specimens (hereafter referred to as "specimens") that is conducted or supported by HHS. This document does the following:

1. Provides guidance as to when research involving coded private information or specimens is or is not research involving human subjects, as defined under HHS regulations for the protection of human research subjects (45 CFR part 46).

2. Reaffirms OHRP policy (see OHRP guidance on *repository activities*http://www.hhs.gov/ohrp/policy/reposit.html and *research on human embryonic stem cells*http://www.hhs.gov/ohrp/archive/references/HESCGuidance.pdf) that, under certain limited conditions, research involving **only** coded private information or specimens is not human subjects research.

3. Clarifies the distinction between (a) research involving coded private information or specimens that does not involve human subjects and (b) human subjects research that is exempt from the requirements of the HHS regulations.

4. (4) References pertinent requirements of the HIPAA Privacy Rule that may be applicable to research involving coded private information or specimens.

NOTE: Some HHS conducted or supported research involving coded private information or specimens may be subject to Food and Drug Administration (FDA) regulations. The FDA regulatory definitions of human subject (21 CFR 50.3(g), 21 CFR 56.102(e)) and subject (21 CFR 312.3(b), 21 CFR 812.3 (p)) differ from the definition of human subject under HHS regulations at 45 CFR 46.102(f). This guidance document does not apply to research regulated by FDA that involves coded private information or specimens. Anyone needing guidance on such FDA-regulated research should contact the FDA.

Target Audience: Institutional review boards (IRBs), investigators, and funding agencies that may be responsible for review or oversight of human subjects research conducted or supported by HHS.

Background:

HHS regulations define *research* at 45 CFR 46.102(d) as follows:

> *Research* means a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge. Activities which meet this definition constitute research for purposes of this policy, whether or not they are conducted or supported under a program which is considered research for other purposes. For example, some demonstration and service programs may include research activities.

HHS regulations define *human subject* at 45 CFR 46.102(f) as follows:

Human subject means a living individual about whom an investigator (whether professional or student) conducting research obtains

1. data through intervention or interaction with the individual, or
2. identifiable private information.

Intervention includes both physical procedures by which data are gathered (for example, venipuncture) and manipulations of the subject or the subject's environment that are performed for research purposes. Interaction includes communication or interpersonal contact between investigator and subject. *Private information* includes information about behavior that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place, and information which has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public (for example, a medical record). 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 for obtaining the information to constitute research involving human subjects (bolding added for emphasis).

For purposes of this document, *coded* means that:

1. identifying information (such as name or social security number) that would enable the investigator to readily ascertain the identity of the individual to whom the private information or specimens pertain 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.

OHRP considers the term *investigator* to include anyone involved in conducting the research. OHRP does not consider the act of solely providing coded private information or specimens (for example, by a tissue repository) to constitute involvement in the conduct of the research. Note that if the individuals who provide coded information or specimens collaborate on other activities related to the conduct of this research with the investigators who receive such information or specimens, then OHRP would consider such additional activities to constitute involvement in the conduct of the research. Examples of such additional activities include, but are not limited to: (1) the study, interpretation, or analysis of the data resulting from the coded information or specimens; and (2) authorship of presentations or manuscripts related to the research.

Guidance:

Under the definition of human subject at 45 CFR 46.102(f), *obtaining* identifiable private information or identifiable specimens for research purposes constitutes human subjects research. *Obtaining* identifiable private information or identifiable specimens includes, but is not limited to:

1. using, studying, or analyzing for research purposes identifiable private information or identifiable specimens that have been provided to investigators from any source; and
2. using, studying, or analyzing for research purposes identifiable private information or identifiable specimens that were already in the possession of the investigator.

In general, OHRP considers private information or specimens to be individually identifiable as defined at 45 CFR 46.102(f) when they can be linked to specific individuals by the investigator(s) either directly or indirectly through coding systems.

Conversely, OHRP considers private information or specimens not to be individually identifiable when they cannot be linked to specific individuals by the investigator(s) either directly or indirectly through coding systems. For example, OHRP does not consider research involving **only** coded private information or specimens to involve human subjects as defined under 45 CFR 46.102(f) if the following conditions are both met:

1. the private information or specimens were not collected specifically for the currently proposed research project through an interaction or intervention with living individuals; and
2. the investigator(s) cannot readily ascertain the identity of the individual(s) to whom the coded private information or specimens pertain because, for example:
    a. 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 (note that the HHS regulations do not require the IRB to review and approve this agreement);

b. 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 investigators under any circumstances, until the individuals are deceased; or

c. there are other legal requirements prohibiting the release of the key to the investigators, until the individuals are deceased.

This guidance applies to existing private information and specimens, as well as to private information and specimens to be collected in the future for purposes other than the currently proposed research. The following are examples of private information or specimens that will be collected in the future for purposes other than the currently proposed research: (1) medical records; and (2) ongoing collection of specimens for a tissue repository.

In some cases an investigator who obtains coded private information or specimens about living individuals under one of the conditions cited in 2(a)-(c) above may (1) unexpectedly learn the identity of one or more living individuals, or (2) for previously unforseen reasons now believe that it is important to identify the individual(s). If, as a result, the investigator knows, or may be able to readily ascertain, the identity of the individuals to whom the previously obtained private information or specimens pertain, then the research activity now would involve human subjects under the HHS regulations. Unless this human subjects research is determined to be exempt under HHS regulations at 45 CFR 46.101(b), IRB review of the research would be required. Informed consent of the subjects also would be required unless the IRB approved a waiver of informed consent under HHS regulations at 45 CFR part 46.116(c) or (d).

Who Should Determine Whether Human Subjects are Involved in Research

OHRP recommends that institutions have policies in place that designate the individual or entity authorized to determine whether research involving coded private information or specimens constitutes human subjects research. The person(s) authorized to make the determination should be knowledgeable about the human subject protection regulations. In addition, the institution should ensure the appropriate communication of such a policy to all investigators. OHRP recommends that investigators not be given the authority to make an independent determination that research involving coded private information or specimens does not involve human subjects.

Research not Involving Human Subjects Versus Exempt Human Subjects Research

OHRP is aware that questions often are raised regarding the distinction between research involving private information or specimens that does not involve human subjects (as above) and human subjects research that is exempt from the requirements of HHS regulations at 45 CFR part 46. This distinction can be made easier by always using the following sequential assessment when evaluating a particular activity conducted or supported by HHS:

1. Does the activity involve *research*? If yes, proceed to question (2). If no, 45 CFR part 46 does not apply to the activity.

2. Does the activity involve *human subjects*? If yes, proceed to question (3). If no, 45 CFR part 46 does not apply to the activity.

   In analyzing a particular activity under the second question, it is important to focus on what is being **obtained** by the investigators. If the investigators are not obtaining either data through intervention or interaction with living individuals, or identifiable private information, then the research activity does not involve human subjects. Therefore, no assessment of the research activity using the third question below regarding exemptions is required because the exemptions provided for under 45 CFR 46.101(b) apply only to research involving human subjects.

3. Is the activity exempt under HHS regulations at 45 CFR 46.101(b)? If yes, 45 CFR part 46 does not apply. If no, 45 CFR part 46 does apply.

   With respect to research involving private information and specimens, the exemption that is most frequently relevant is the exemption under HHS regulations at 45 CFR 46.101(b)(4):

   "Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects."

Having determined under the second question above that a research activity involves human subjects because the investigators are obtaining identifiable private information or specimens, assessment under the exemption at 45 CFR 46.101(b)(4) focuses, in part, on: (1) whether the data or specimens are **existing** at the time the research is proposed to an institutional official or IRB for a determination of whether the research is

exempt, and (2) how the data or information is recorded by the investigators. This exemption would not apply if the investigators, having obtained identifiable private information or specimens from existing records or specimens, record the data or information in a coded manner, since the code would enable subjects to be identified through identifiers linked to the subjects.

To demonstrate how the determination of whether a research study is human subjects research differs from the determination of whether a human subjects research study is exempt under 45 CFR 46.101(b)(4), consider the following examples, in which an investigator obtains health information of living patients who were treated for arthritis with either Drug A or Drug B. The investigator obtains this information in order to evaluate and compare the treatment outcomes associated with these two drugs:

1. An investigator obtains only coded information on the treatment outcomes of patients treated for arthritis with Drug A versus Drug B from the patients' treating physician. The only involvement of the treating physician is to provide coded information to the investigator. The investigator and the treating physician enter into an agreement prohibiting the release of the key to decipher the code to the investigator under any circumstances, until the individuals are deceased. In this example, the investigator is not conducting human subjects research because the investigator cannot readily ascertain the patients' identity.

2. An investigator obtains individually identifiable information on the treatment outcomes of patients treated for arthritis with either Drug A or Drug B by viewing patients' existing individually identifiable medical records at the clinics where the patients were treated. The investigator records the patients' treatment outcomes in a coded manner that could permit the identification of the patients. In this example, the investigator is conducting human subjects research because the investigator is obtaining identifiable private information from patients' (and now subjects') medical records. The study would not be exempt under 45 CFR 46.101(b)(4) since the investigator is recording the information in a coded manner, thus allowing the subjects to be identified indirectly through identifiers linked to the subjects.

3. An investigator obtains individually identifiable information on the treatment outcomes of patients treated for arthritis with either Drug A or Drug B by viewing patients' existing individually identifiable medical records at the clinics where the patients were treated. The investigator records only patient age, sex, diagnosis, treatment, and health status at the end of 6 months of treatment so that the investigator cannot link the recorded information back to the patients. In this example, the investigator is conducting human subjects research because the investigator is obtaining identifiable private information from patients' (and now subjects') medical records. However, the study would be exempt under 45 CFR 46.101(b)(4) since the investigator records the information in such a manner that subjects cannot be identified either directly or indirectly through identifiers linked to the subjects.

Comparison to the HIPAA Privacy Rule

The Privacy Rule is a Federal regulation under the Health Insurance Portability and Accountability Act (HIPAA) of 1996 (see 45 CFR part 160 and subparts A and E of part 164). The Privacy Rule permits covered entities under the Rule to determine that health information is de-identified even if the health information has been assigned, and retains, a code or other means of record identification, provided that:

1. the code is not derived from or related to the information about the individual;

2. the code could not be translated to identify the individual; and

3. the covered entity under the Privacy Rule does not use or disclose the code for other purposes or disclose the mechanism for re-identification (see HHS guidance entitled, *Institutional Review Boards and the HIPAA Privacy Rule*, page 6, Q and A #3, at http://privacyruleandresearch.nih.gov/pdf/IRB_Factsheet.pdf).

Regarding condition (1) above, in contrast to the Privacy Rule, information that is linked with a code derived from identifying information or related to information about the individual is not considered to be individually identifiable under the HHS regulations for the protection of human subjects at 45 CFR 46.102(f), if the investigators cannot readily ascertain the identity of the individual(s) to whom the coded private information or specimen pertains. Therefore, some coded information, in which the code has been derived from identifying information linked to or related to the individual, would be individually identifiable under the Privacy Rule, but might not be individually identifiable under 45 CFR part 46.

If you have specific questions about how to apply this guidance, please contact OHRP by phone at (866) 447-4777 (toll-free within the U.S.), (240) 453-6900, or by e-mail at ohrp@hhs.gov.

Documents in PDF format require the Adobe Acrobat Reader®. If you experience problems with PDF documents, please download the latest version of the Reader®.

Research Involving Private Information or Biological Specimens

Are the specimens/data obtained from living individuals?

- NO, individuals are NOT living → NOT Human Subjects Research
- YES, individuals ARE living → Are the specimens/data:
  - Human cell lines obtained from a commercial provider (e.g. ATCC); or
  - Human cells about which all information has been published; or
  - Unidentifiable specimens/data obtained from a commercial provider; or
  - Unidentifiable specimens/data obtained from an provider that is prohibited from releasing identifiers by established regulations or policies

  - YES → NOT Human Subjects Research
  - NO → Were/will the specimens/data (be) collected specifically for the proposed research through an interaction or intervention with living individuals?

    - YES → Human Subjects Research
    - NO → Can the *recipient* link the specimens/data **directly** to identifiable living individuals?

      - YES → Human Subjects Research
      - NO → Can the *provider* link the specimens/data, directly or indirectly, to identifiable living individuals?

        - NO → NOT Human Subjects Research
        - YES → Does the *provider* meet the definition of an "investigator" in the recipient's research?

          - YES, *provider* is collaborating in recipient's research → Human Subjects Research
          - NO, *provider* is "solely providing" → Are the specimens/data provided with a code linking them to identifiable living individuals?

            - NO → NOT Human Subjects Research
            - YES → Can the *recipient* readily ascertain the identities of the individuals to whom the specimens/data pertain? Examples of situations in which the recipient **cannot** link the specimens/data to living individuals include:
              - the key to decipher the code is destroyed before the research begins; or
              - the investigators and the holder of the key to the code enter into an agreement preventing the release of the key to investigators under any circumstances; or
              - there are IRB-approved written policies in place preventing the release of the key under any circumstances; or
              - there are other legal requirements prohibiting the release of the key under any circumstances.

              - NO → NOT Human Subjects Research
              - YES → Human Subjects Research

FINAL
OEP/OER
June 6, 2005

Frequently Asked Questions

Public FAQs
NIH Staff FAQs

Human Subjects Research - Human Specimens, Cell Lines or Data

Last Revised: February 11, 2010

Related NIH Staff FAQs

Back to Top

1. **When does research with human specimens, cells, cell lines, or data involve human subjects?**
In order for research with human specimens, cells, cell lines, or data to involve human subjects,

   1. The specimens, cells, or data:

      - Must be or must have been obtained from individuals who are alive; AND
      - Must be or must have been obtained by an investigator conducting research

   AND

   2. The investigator EITHER:

      - Must be obtaining or must have obtained specimens, cells, or data through interaction or intervention with living individuals; OR
      - Must be obtaining or have obtained individually identifiable private information.

   IF providers of **coded** human specimen, cells, cell lines or data:

      - Obtained or will obtain the specimens or data, AND
      - Can link the specimens or data to living individuals, AND
      - Will also collaborate on other activities related to the conduct of a proposed research project with the investigators who obtain the specimens or data;

   THEN **both** the providers and recipients will be considered to be involved in the conduct of the research and are conducting human subjects research.

2. **What are examples of research involving human specimens, cells, cell lines, or data that would not be considered human subjects research under HHS regulations at 45 CFR Part 46?**
Research that proposes the use of only cadaver specimens is not human subjects research, because, by definition, human subjects must be "living individuals." Research involving cadaver specimens, therefore, is not regulated by 45 CFR Part 46, but may be governed by other Federal, state and local laws.

   Research that proposes the use of human cell lines available from the American Type Culture Collection or a similar repository is not considered human subjects research because the cells are publicly available and all of the information known about the cell lines (perhaps, including the donor) is also publicly available.

   Research that proposes the use of established cells from a donor whose identity cannot be readily ascertained by the investigator is not considered to be human subjects research, either, for example, because there are written IRB-approved policies and procedures prohibiting release of identifiers or an agreement specifying that identifying information will not be shared with the investigator.

3. **What are examples of research involving coded human specimens, cells, cell lines, or data that is not human subjects research?**
Research that involves **only coded** human biological specimens or coded private information/data from living individuals is not human subjects research under the HHS human subjects regulations (45 CFR Part 46) if:

(1) the specimens, cells, cell lines or private information/data were not collected specifically for the currently proposed research project through an interaction or intervention with living individuals;
and
(2) the investigator(s) cannot readily ascertain the identity of the individual(s) to whom the coded private information or specimens pertain because, for example:
(a) the key to decipher the code is destroyed before the research begins;
(b) 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 (note that the HHS regulations do not require the IRB to review and approve this agreement);
(c) 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 investigators under any circumstances, until the individuals are deceased; or
(d) there are other legal requirements prohibiting the release of the key to the investigators, until the individuals are deceased.

Refer to the following guidance from the Office for Human Research Protections (OHRP) for additional information and examples: http://www.hhs.gov/ohrp/policy/cdebiol.pdf.

4. **Am I proposing human subjects research if my studies will use ONLY cell lines?**
Research that proposes the use of human cell lines available from the American Type Culture Collection or a similar repository is not considered human subjects research because the cells are publicly available and all of the information known about the cell lines (perhaps, including the donor) is also publicly available.

Research that proposes the use of established cells from a donor whose identity cannot be readily ascertained by the investigator is not considered to be human subjects research, either for example, because there are written policies and procedures prohibiting release of identifiers and/or an agreement specifying that identifying information will not be shared with the investigator.

Research with Primary cells: If you are taking blood or other cells from living individuals for research purposes in order to create a cell line, the research involves human subjects.

5. **Am I proposing human subjects research if I obtain specimens/data from a repository or database?**
IF the repository/database obtains the specimens/data without identifiers, **OR**

IF the repository/database obtains the specimens/data with identifiers but is prevented, by law*, from providing identifiers that link to living individuals and the repository/database plays no collaborative role in the proposed research,

THEN studies using the specimens/data from the repository/database are not human subjects research.

If your proposed studies meet these criteria, you should indicate "No" for human subjects, and you should consider including an explanation of why your research does not involve human subjects in the Human Subjects section in your application

If your proposed studies do not meet these criteria, you are proposing human subjects research. You should indicate "Yes" for human subjects and complete the Human Subjects section in your application.

*Because NIH is not a covered entity, NIH cannot comment on the adequacy with which institutions address HIPAA requirements. NIH, therefore, cannot recognize the HIPAA Privacy Act as the law preventing the release of identifiers.

6. **What are the requirements for Exemption 4 (E4)?**
HHS regulations state:

"Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects." 46.101(b)(4)

If you receive or have access to existing individually identifiable private information or identifiable specimens from living individuals (e.g., pathology or medical records), you are conducting human subjects research. If you as the investigator or your collaborator record the information in such a manner that you cannot subsequently access or obtain direct or indirect identifiers that are linked to the subjects, research activities that involve data recorded in this manner meets the requirements of Exemption 4. If you will retain or can access any identifiers, the research project is not exempt under Exemption 4.

7. **Is research that meets the criteria for Exemption 4 considered human subjects research?**
Yes. Research that meets the criteria for Exemption 4 is Human Subjects Research.

Exemption 4 includes research involving the collection or study of **existing** data, documents, records, pathological specimens, or diagnostic specimens, if these sources are **publicly available** or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through **identifiers** linked to the subjects.

Please note: human subjects research that meets the criteria for Exemption 4 is not considered "clinical research" as defined by NIH; therefore, the NIH policies for addressing inclusion of women, minorities and children do not apply to research that is determined to meet the criteria for Exemption 4.

8. **How can I determine whether my research meets the criteria for Exemption 4?**
The humans subjects regulations decision charts from the Office of Human Research Protection (OHRP) will help you to see whether your research falls under the human subjects regulations and if so, whether it meets the criteria for Exemption 4.

Please note: OHRP advises that investigators should not have the authority to make an independent determination that research involving human subjects is exempt. OHRP guidance states that Exemptions should be independently determined. Institutions often designate their IRB to make this determination. Because NIH does not require IRB approval at time of application, the exemptions designated often represent the opinion of the Principal Investigator, and the justification(s) provided by the Principal Investigator for the exemption(s) is/are evaluated during peer review.

9. **What is meant by "existing" data or specimens?**
Exemption 4 applies to retrospective studies of specimens and/or data that have already been collected. The materials must be "on the shelf" (or in the freezer) at the time the protocol is submitted to the IRB or other designated officials at your institution to

determine whether the research is indeed exempt. Research that involves the ongoing collection of specimens and/or data does not meet the criteria for Exemption 4, even if they were destined to be discarded.

10. **What is meant by "publicly available sources"?**
    This language in the regulation was intended to apply to public sources of data, such as census data. Its meaning with respect to human tissue specimens is widely debated. Although there are organizations that make human cells and tissues broadly accessible to the research community, these materials are not usually available to the public at large and are not generally considered to be publicly available.

11. **What is meant by "identifiers linked to the subjects"?**
    Examples of identifiers, would include names, social security numbers, medical record numbers, or pathology accession numbers, or any other "code" that permits specimens or data to be linked to individually identifiable living individuals and perhaps also to associated medical information.

Back to Top

This page last updated on February 11, 2010
Content Manager: OEP-HS@mail.nih.gov
Technical Issues: E-mail OER Webmaster