Coded Private Information or Biospecimens Used in Research, Guidance (2018)

NOTE: THIS GUIDANCE REPLACES OHRP'S 2008 GUIDANCE ENTITLED "CODED PRIVATE INFORMATION OR SPECIMENS USED IN RESEARCH." THIS GUIDANCE HAS BEEN UPDATED TO BE CONSISTENT WITH THE COMMON RULE ISSUED JANUARY 19, 2017 AND EFFECTIVE JANUARY 19, 2018.

CODED PRIVATE INFORMATION OR BIOSPECIMENS USED IN RESEARCH

This guidance represents the Office for Human Research Protection's (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 Health and Human Services (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: January 19, 2018

Scope: This document provides guidance as to when research involving coded private information or coded biospecimens involves a human subject, as defined under HHS regulations for the protection of human research subjects (45 CFR 46.102(e)). Specifically, this guidance does the following:

1. Clarifies that research in which an investigator obtains information or biospecimens through intervention or interaction with the individual, and then uses, studies, or analyzes the information or biospecimen for research purposes is research that involves human subjects, even if that investigator has coded the information or biospecimens.

2. Clarifies that, under certain limited conditions, secondary research involving **only** coded private information or coded biospecimens is research that does not involve human subjects

3. References pertinent requirements of the Health Insurance Portability and Accountability Act's (HIPAA's) Privacy Rule that may be applicable to research involving coded private information or coded biospecimens.

NOTE: Some HHS-conducted or -supported research involving coded private information or coded biospecimens 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(e). This guidance document does not apply to research regulated by FDA that involves coded private information or coded biospecimens. Anyone needing guidance on such FDA-regulated research should contact FDA.

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

Background:

HHS regulations define the term "research" at 45 CFR 46.102(l) as follows:

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

The following activities are deemed not to be research, even if they involve the collection and analysis of coded private information or coded biospecimens:

(1) Scholarly and journalistic activities (e.g., oral history, journalism, biography, literary criticism, legal research, and historical scholarship), including the collection and use of information, that focus directly on the specific individuals about whom the information is collected.

(2) Public health surveillance activities, including the collection and testing of information or biospecimens, conducted, supported, requested, ordered, required, or authorized by a public health authority. Such activities are limited to those necessary to allow a public health authority to identify, monitor, assess, or investigate potential public health signals, onsets of disease outbreaks, or conditions of public health importance (including trends, signals, risk factors, patterns in diseases, or increases in injuries from using consumer products). Such activities include those associated with providing timely situational awareness and priority setting during the course of an event or crisis that threatens public health (including natural or man-made disasters).

(3) Collection and analysis of information, biospecimens, or records by or for a criminal justice agency for activities authorized by law or court order solely for criminal justice or criminal investigative purposes.

(4) Authorized operational activities (as determined by each agency) in support of intelligence, homeland security, defense, or other national security missions.

HHS regulations define the term “human subject” at 45 CFR 46.102(e) as follows:

Human subject means a living individual about whom an investigator (whether professional or student) 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**.**

The regulations state that the term “intervention” includes physical procedures by which information or biospecimens are gathered (e.g., venipuncture) and manipulations of the subject or the subject’s environment that are performed for research purposes. The term “interaction” includes communication or interpersonal contact between investigator and subject. With regard to (2) above:

- 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 that has been provided for specific purposes by an individual and that the individual can reasonably expect will not be made public (e.g., a medical record).

- Identifiable private information is private information for which the identity of the subject is or may readily be ascertained by the investigator or is associated with the information.

- An identifiable biospecimen is a biospecimen for which the identity of the subject is or may readily be ascertained by the investigator or is associated with the biospecimen.

For the purposes of this document, the term “secondary research” refers to reusing information and biospecimens that are collected for some other primary or initial activity. The information collected could be collected either for research studies other than the proposed research or for nonresearch purposes. The information or biospecimens do not have to be existing at the time that the secondary study is proposed. Such information or biospecimens could be: (1) found by the investigator in some type of records, archive, database, or on the Internet (in the case of information) or some type of tissue repository (such as a hospital’s department for storing clinical pathology specimens); (2) collected simultaneously at the time of a clinical or other type of service or intervention; or (3) stored in a research facility for the purposes of another research study. Secondary research does not include obtaining information or biospecimens through a research intervention or interaction with an individual.

For purposes of this document, the term “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 biospecimens 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 biospecimens.

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 biospecimens (e.g., by a tissue repository) to constitute involvement in the conduct of the research. Note that if the individuals who provide coded information or coded biospecimens collaborate on other activities related to the conduct of this research with the

investigators who receive such information or biospecimens, then OHRP would consider these additional activities to constitute human subjects 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 coded biospecimens; and (2) authorship of presentations or manuscripts related to the research.

Guidance:

Under the definition of human subject at 45 CFR 46.102(e)(1)(i), a research study involves human subjects if an investigator obtains information or biospecimens through intervention or interaction with the individual, and uses, studies, or analyzes the information or biospecimens. It is the case that such research involves human subjects even if the investigator codes the information or biospecimens, or removes all direct or indirect identifiers before using, studying, or analyzing the information or biospecimens. This is because the investigator is also the person rendering the biospecimens or information coded or nonidentifiable for research purposes, and the identity of the individuals from whom the information or biospecimens were obtained might be readily ascertainable to the investigator. Note that in this circumstance, the activity might be eligible for an exemption (see 45 CFR 46.104(d)(4)).

If the human subjects research is not exempt under the regulations, and there is more than one institution involved in the conduct of the study, it might be the case that an institution using, studying, or analyzing the coded information or biospecimens would not be considered to be engaged in the research. Institutions not considered to be engaged in research: (1) need not hold or obtain an OHRP-approved Federalwide Assurance (FWA) (45 CFR 46.103(a)), or (2) certify to the HHS agency conducting or supporting the research that the research has been reviewed and approved by an IRB. For OHRP's guidance on the Engagement of Institutions in Human Subjects Research, see https://www.hhs.gov/ohrp/regulations-and-policy/guidance/guidance-on-engagement-of-institutions/index.html) (https://www.hhs.gov/ohrp/regulations-and-policy/guidance/guidance-on-engagement-of-institutions/index.html) . (The scenario at B.(7) in OHRP's engagement guidance discusses when institutions whose employees or agents obtain coded private information or coded biospecimens from another institution involved in the research are considered to be not engaged in human subjects research.)

Under the definition of human subject at 45 CFR 46.102(e)(1)(ii), the criterion for determining whether secondary research involves human subjects is whether an investigator obtains, uses, studies, analyzes, or generates identifiable private information or identifiable biospecimens for research purposes. Obtaining identifiable private information or identifiable biospecimens includes:

1. using, studying, or analyzing for research purposes identifiable private information or identifiable biospecimens that have been provided to investigators from any source; and

2. using, studying, analyzing, or generating for research purposes identifiable private information or identifiable biospecimens that were already in the possession of the investigator.

In general, OHRP considers private information or biospecimens to be individually identifiable as defined at 45 CFR 46.102(e) 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 biospecimens not to be individually identifiable when they cannot be linked to specific individuals by the investigator(s) either directly or indirectly through coding systems, or their identities cannot readily be associated by the investigator through some other means. For example, OHRP does not consider secondary research involving **only** coded private information or coded biospecimens to involve human subjects as defined under 45 CFR 46.102(e) if the following condition is met:

The investigator(s) cannot readily ascertain the identity of the individual(s) to whom the coded private information or coded biospecimens pertain because, for example:

1. 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);

2. 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

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

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

In some cases, an investigator who obtains coded private information about or coded biospecimens from living individuals under one of the conditions cited in examples 1-3 above may unexpectedly learn the identity of one or more living individuals, or for previously unforeseen 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 biospecimens 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 the HHS regulations at 45 CFR 46.104(d), IRB review of the research would be required. (Note that some of the exemptions at 45 CFR 46.104(d) require limited IRB review.)

For all nonexempt research, informed consent of the subjects also would be required unless the IRB approved a waiver of informed consent under HHS regulations at 45 CFR 46.116(e) or (f). (Note that the exemptions at 45 CFR 46.104(d)(7) and (8) regarding the storage, maintenance, and secondary research use of identifiable private information or identifiable biospecimens require that broad consent be obtained from subjects under 45 CFR 46.116(d)).

Who Should Determine Whether Human Subjects are Involved in Secondary Research?

OHRP recommends that institutions have policies in place that designate the individual or entity authorized to determine whether secondary research involving coded private information or coded biospecimens constitutes nonexempt or exempt 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 coded biospecimens does not involve human subjects.

Secondary Research Not Involving Human Subjects Versus Exempt Human Subjects Research

OHRP is aware that questions are raised regarding the distinction between secondary research involving private information or biospecimens 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 yes, it is also possible that one of the secondary research exemptions at 45 CFR 46.104(d)(4), (7), or (8) might apply. If no, 45 CFR part 46 does not apply to the activity.

3. Is the activity exempt under HHS regulations at 45 CFR 46.104? If yes, some aspects of 45 CFR part 46 still might apply, for example, the need for limited IRB review or obtaining broad consent. If no, 45 CFR part 46 does apply.

In analyzing a particular secondary research activity under the second question, it is important to focus on what is being **obtained** by the investigators. If none of the investigators conducting the research study are obtaining either data through intervention or interaction with living individuals, or identifiable private information or identifiable biospecimens, 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 apply only to research involving human subjects. The exemptions provided under 45 CFR 46.104(d)(4), (7), and (8) are the exemptions that may be relevant for secondary research involving human subjects when using identifiable private information or identifiable biospecimens.

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* at http://privacyruleandresearch.nih.gov/pdf/IRB_Factsheet.pdf - PDF - PDF (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(e), if the investigators do not know and cannot readily ascertain the identity of the individual(s) to whom the coded private information or

coded biospecimen pertains. Therefore, some coded information, or coded biospecimens 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 United States), (240) 453-6900, or by e-mail at ohrp@hhs.gov (mailto:ohrp@hhs.gov) .

OHRP Headquarters

Office for Human Research Protections
1101 Wootton Parkway, Suite 200
Rockville, MD 20852