UNIVERSITY AT BUFFALO
HUMAN RESEARCH PROTECTIONS PROGRAM

**Certification of De-Identification
of Health Information for Research Purposes**

INSTRUCTIONS

Health information is considered "de-identified" under the HIPAA privacy regulations if specific categories of direct or indirect personal identifiers are removed from data sets. The categories of identifiers are listed on the enclosed certification form. De-identified information is considered anonymous and does not fall under any of the regulatory provisions of HIPAA.

When information is not being acquired from a HIPAA covered entity, the research may be considered de-identified as long as none of the personal identifiers listed are maintained by the research team.

If information is being provided to a researcher by an entity that is covered by HIPAA, the information from that providing entity may be considered de-identified as long as none of the listed identifiers are _ever_ disclosed to the research team (including the PI) by any means, including the formal transfer of identifiers to the researcher, or an incidental disclosure of the identifiers, such as in a conversation with covered entity staff about the research project.

A limited mechanism does exist for acquiring de-identified data from a HIPAA covered entity while maintaining a way to re-establish identity if necessary or required by law. In these instances the covered entity may disclose a unique ID linked to the research subject along with their associated de-identified data to the researcher. Any such key must be generated by the covered entity, must not be based on any personal identifiers, and the algorithm for generating the key or re-linking it with the subjects identity may not be disclosed if it would permit the researcher any means to independently re-identify the individual associated with the key. While this mechanism does not permit the researcher to re-identify the research subject directly, it does permit re-identification by the HIPAA covered entity when necessary.

Permissible Identifiers: Health information is "de-identified" even if the following identifiers are used:

- gender,
- ethnicity,
- age (if less than 90)
- year but not dates of birth, admission, discharge, death, etc.
- the first three digits of most zip codes (exceptions noted under (B). below).

UNIVERSITY AT BUFFALO
HUMAN RESEARCH PROTECTIONS PROGRAM

Certification of De-Identification of
Health Information for Research Purposes

Investigator

Protocol Title

Please **check off** each of the categories below to verify that these classes of identifiers pertaining to the research subject, their relatives, employers, or household members, can **not** be linked with the research subject's health information generated or acquired as part of this protocol. These elements must be removed from health information, and the below affirmation provided, in order to consider health information de-identified per CFR 164.514(b)(2):

_____ (A) Names;

_____ (B) All geographic subdivisions smaller than a State, including street address, city, county, precinct, zip code, and their equivalent geocodes, except for the initial three digits of a zip code if, according to the current publicly available data from the Bureau of the Census: (*1*) The geographic unit formed by combining all zip codes with the same three initial digits contains more than 20,000 people; and (*2*) The initial three digits of a zip code for all such geographic units containing 20,000 or fewer people is changed to 000.

_____ (C) All elements of dates (except year) for dates directly 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) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older;

_____ (D) Telephone numbers;

_____ (E) Fax numbers;

_____ (F) Electronic mail addresses;

_____ (G) Social security numbers;

_____ (H) Medical record numbers;

_____ (I) Health plan beneficiary numbers;

_____ (J) Account numbers;

_____ (K) Certificate/license numbers;

_____ (L) Vehicle identifiers and serial numbers, including license plate numbers;

_____ (M) Device identifiers and serial numbers;

_____ (N) Web Universal Resource Locators (URLs);

_____ (O) Internet Protocol (IP) address numbers;

_____ (P) Biometric identifiers, including finger and voice prints;

_____ (Q) Full face photographic images and any comparable images; and

_____ (R) Any other unique identifying number, characteristic, or code, except as permitted by paragraph (c) of this section.[1]

I affirm that this research protocol will not collect any of the personal identifiers listed above and that, to the best of my knowledge, no information generated or collected as part of this research protocol can be used alone or in combination with other information in our possession to identify an individual who is the subject of the information.

Principal Investigator signature   Date

---

[1] 164.514(2)(c): A covered entity may assign a code or other means of record identification to allow information de-identified under this section to be re-identified by that covered entity, provided that: (1) *Derivation*. The code or other means of record identification is not derived from or related to information about the individual and is not otherwise capable of being translated so as to identify the individual; and (2) *Security*. The covered entity does not use or disclose the code or other means of record identification for any other purpose, and does not disclose the mechanism for re-identification.