

# De-identification of Protected Health Information and Limited Data Sets

HIPAA P-06



## About This Policy

**Effective:** 7/22/2013

**Last Updated:** 12/13/2021

**Responsible University Office:**

Office of the Chief Privacy Officer

**Responsible University Administrator:**

Chief Privacy Officer

mawerlin@iu.edu

**Policy Contact:**

HIPAA Privacy Officer

HIPAA Security Officer

[hipaa@iu.edu](mailto:hipaa@iu.edu)

## Scope

This policy applies to the workforce members in the designated Indiana University (IU) HIPAA Covered Healthcare Components and HIPAA Affected Areas, anyone rendering services as a Business Associate, and anyone who creates, receives, maintains, or transmits Protected Health Information (PHI) in any capacity at IU, including, but not limited to, faculty, staff, students, trainees, volunteers, visiting scholars, and third-party agents. For the purposes of this policy, all of the above will be referred to as workforce members.

## Policy Statement

Indiana University respects the privacy of all members of the IU community and strives to implement measures to protect privacy consistent with the university mission and environment, applicable legal requirements and professional standards, generally accepted privacy norms, and available resources.

While HIPAA imposes many restrictions on the use and disclosure of protected health information, HIPAA does not regulate the use or disclosure of de-identified information and imposes lesser restrictions on the use and disclosure of Limited Data Sets. It is therefore the policy of Indiana University to use and/or disclose de-identified information or Limited Data Sets where appropriate, in accordance with the policy set forth below. De-identified information and/or Limited Data Sets may still be subject to other confidentiality requirements (e.g., because the information is proprietary) and should be marked confidential when appropriate.

## Reason for the Policy

This policy has two purposes, which are as follows:

1. To specify the requirements for de-identifying Protected Health Information (PHI) in accordance with the HIPAA regulations so that the information will no longer be considered PHI and no longer subject to HIPAA.
2. To specify the requirements for removing certain identifying information from PHI in order to create a Limited Data Set that may be disclosed for research, public health, or health care operations purposes

once the recipient of the PHI enters into a Data Use Agreement. Data in the form of a Limited Data Set is still considered PHI and protected under HIPAA.

## Procedures

### I. DE-IDENTIFIED INFORMATION POLICY:

Health Information is not subject to the HIPAA Privacy Rule if it is de-identified in accordance with the HIPAA Privacy Rule. Individual authorization is not required to use or disclose health information that is de-identified. Health information is considered de-identified if: (a) it does not identify an individual; and (b) there is no reasonable basis to believe it can be used to identify an individual.

### A. Methods for De-Identification

A covered entity may determine that health information is not individually identifiable health information only if one of two methods is used to de-identify health information; the Safe Harbor method of removing identifiers or expert determination.

1. Method 1 – Removing Identifiers: Removal of all of the following identifiers as they pertain to the Individual or to his/her relatives, employers or household members (collectively referred to below as “Individuals”):
  - a. Names.
  - b. All geographic subdivisions smaller than a State, including:
    - (1) Street address of P.O. Box Number
    - (2) City
    - (3) County
    - (4) Precinct
    - (5) Town
    - (6) Zip codes and their equivalent geocodes, except for the initial three digits of a zip code if, according to current publicly-available data from the Bureau of the Census:
      - i. The geographic unit formed by combining all zip codes with the same three initial digits contains more than 20,000 people; and
      - ii. The initial three digits of a zip code for all such geographic units containing 20,000 or fewer people are changed to 000
  - c. All elements of dates (except year) for dates related to an individual including: dates of birth and death, and admission and discharge dates and all ages over 89, in which case the Individuals’ ages must be categorized as 90 or older.
  - d. Telephone numbers
  - e. Fax numbers
  - f. E-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 number
  - 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.
  - r. Any other unique identifying number, characteristic or code

### **Exception:**

- Any code used to replace the identifiers in metadata or any associated documents cannot be derived from any information related to the individual and the master codes, nor can the method to derive the codes be disclosed. For example, a subject's initials cannot be used to code their data because the initials are derived from their name.
- Additionally, there may not be actual knowledge that the individual could be re-identified from the remaining identifiers in the PHI present. In other words, the information would still be considered identifiable if there was a way to identify the individual even though all of the 18 identifiers were removed.

2. Method 2 – Expert Determination: A person with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable.
  - a. Applying such principles and methods, determines that the risk is very small that the information could be used, alone or in combination with other reasonably available information, by an anticipated recipient to identify an individual who is a subject of the information; and
  - b. Documents the methods and results of the analysis that justify such determination.

When expert determination is required to de-identify PHI, the individual or entity providing the service must be approved by the IU Office of the Chief Privacy Officer. See section III below.

## **II. LIMITED DATA SET POLICY**

A covered entity can use and disclose information in the form of a Limited Data Set without the individual's authorization for purposes of research, public health or healthcare operations if the data are released in conjunction with a Data Use Agreement.

### **A. Limited Data Set:**

A limited data set is information from which direct identifiers have been removed. Specifically, as it relates to the individual or his or her relative, employers or household members, all of the following identifiers must be removed in order for health information to be a "Limited Data Set":

1. Names.
2. Street addresses or RR numbers (other than town, city, state and zip code)
3. Telephone numbers
4. Fax numbers
5. E-mail addresses
6. Social Security numbers
7. Medical record numbers
8. Health plan beneficiary numbers
9. Account numbers
10. Certificate/license numbers
11. Vehicle identifiers and serial numbers, including license plate number
12. Device identifiers and serial numbers
13. Web Universal Resource Locators (URLs)
14. Internet Protocol (IP) address numbers
15. Biometric identifiers (including finger and voice prints).
16. Full-face photographic images.

**B. Health information that may remain in the information disclosed includes:**

1. Dates such as date of birth, date of death, admission, discharge, service;
2. City, state, five-digit zip code;
3. Ages in years, months or days or hours; and
4. Unique identifying numbers, characteristics or codes provided the unique identifiers cannot reasonably be used to identify an individual and provided they are not derived from excluded elements (e.g. initials).

**C. Data Use Agreement:**

The Data Use Agreement must contain the following elements:

1. A description of the permitted uses and disclosures of the Limited Data Set, which must be limited to and consistent with public health, research or health care operations purposes;
2. A description of those persons who are permitted to use or receive the Limited Data Set;
3. A statement requiring that the Limited Data Set recipient will:
  - a. Not use or further disclose the information other than as permitted in the Data Use Agreement or as required by law;
  - b. Use appropriate safeguards to prevent the use or disclosure of the information other than as permitted in the Data Use Agreement;
  - c. Report to the Indiana University any use or disclosure of the information that is not permitted by the Data Use Agreement of which it becomes aware;
  - d. Ensure that any of its agents or subcontractors to whom it provides the Limited Data Set agrees to the same restrictions and conditions that apply to the Limited Data Set recipient; and
  - e. Not identify the information or contact the Individuals who are the subject of the information.

**III. AUTHORIZED INDIVIDUAL TO DE-IDENTIFY DATA OR CREATE LIMITED DATA SETS**

Only Indiana University Workforce or third-party Business Associates with whom Indiana University has contracted may de-identify health information or use the health information to create Limited Data Sets. If a third-party Business Associate is used for this purpose, then there must be a Business Associate Agreement in place with such third-party.

**IV. NON-COMPLIANT LIMITED DATA SET RECIPIENTS**

If at any time Indiana University becomes aware that a recipient of a Limited Data Set has violated his/her/its Data Use Agreement, then Indiana University must:

- A.** Take reasonable steps to end the breach of the agreement or cause the breach to be cured; or
- B.** If the breach cannot be ended or cured, then stop disclosing the Limited Data Set or other PHI to the recipient and report the problem to the Secretary of Health and Human Services.

**Definitions**

**Individually Identifiable Health Information (IIHI):** A subset of health information, including demographic information collected from an individual, and: (1) is created or received by a health care provider, health plan, employer, or health care clearinghouse; and (2) relates to the past, present, or future physical or mental health or condition of an individual; the provision of health care to an individual; or the past, present, or future payment for the provision of health care to an individual; and identifies the individual or there is a reasonable basis to believe the information can be used to identify the individual.

**Minimum Necessary:** A standard that requires covered entities to take reasonable steps to limit the use or disclosure of, and requests for PHI to the minimum necessary to accomplish the intended purpose. The minimum necessary standard does not apply to certain uses or disclosures such as those requests by a health care provider for treatment purposes, disclosures to the individual who is the subject of the information

or pursuant to an individual's authorization.

**Protected Health Information (PHI):** Individually identifiable health information held or transmitted by a covered entity or its business associate in any form or medium, whether electronic, on paper or oral.

See [Glossary of HIPAA Related Terms](#) for complete list of terms.

## History

07/01/2014 Effective Date  
01/13/2016 Updated Definitions Section  
08/01/2016 Updated DUA Template – Appendix 3, added link to Glossary  
06/xx/2017 Published on University policy site  
12/01/2020 Updated Title, Appendices, Related Information, DUA information  
12/13/2021 Updated contact information

## Related Information

HIPAA-A02 General Administrative

HIPAA-G01 HIPAA Sanctions Guidance

HIPAA-G04 Limited Data Set and Data Use Agreement Guidance

HIPAA-P01 Uses & Disclosures of Protected Health Information Policy

HIPAA-P02 Minimum Necessary Policy

HRPP Policy: [Use of Protected Health Information \(PHI\) in Research](#)

HHS De-identification Guidance: *Guidance Regarding Methods for De-Identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule.*  
<http://www.hhs.gov/ocr/privacy/hipaa/understanding/coveredentities/De-identification/guidance.html>

Radiological Society of North America (RSNA) Guidance: *Protecting Patient Information in Medical Presentations, Publications and Products*  
<https://www.rsna.org/-/media/Files/RSNA/Practice-Tools/RemovingPHI.pdf>

IU Office of Research Administration Data Use Agreement Request Form: <https://ora-fireform.eas.iu.edu/online/form/authen/oracr>

IU Knowledge Base Article: Properly redact information at IU  
<https://kb.iu.edu/d/adgl>



Project Name: \_\_\_\_\_

Department: \_\_\_\_\_

IRB Protocol Number, if applicable: \_\_\_\_\_

Contact Name: \_\_\_\_\_

Information is de-identified when none of the following 18 types of identifiers are contained in the information and if no one accessing the information has actual knowledge that the information could be used – alone or in combination with other information – to identify any individual who is the subject of the information.

- |  |   |
|--|---|
| <input type="checkbox"/> Names (including initials)  | <input type="checkbox"/> Fax number   |
| <input type="checkbox"/> Addresses (All geographic subdivisions smaller than a state) <ul style="list-style-type: none"><li>• Street address, city, county, town, precinct</li><li>• Zip code or 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)</li></ul> | <input type="checkbox"/> Email address  |
| <input type="checkbox"/> Dates: All elements of date (except year) directly related to the individual (birth date, death date, admission and discharge dates, clinic visit date) <ul style="list-style-type: none"><li>• All ages over 89 or dates indicating such an age</li></ul>  | <input type="checkbox"/> Social Security Number   |
| <input type="checkbox"/> Telephone number  | <input type="checkbox"/> Medical Record Number  |
|  | <input type="checkbox"/> Health Plan Beneficiary Number   |
|  | <input type="checkbox"/> Account Numbers  |
|  | <input type="checkbox"/> Certificate or license numbers   |
|  | <input type="checkbox"/> Vehicle identification/serial numbers including license plate numbers  |
|  | <input type="checkbox"/> Device identification/serial numbers   |
|  | <input type="checkbox"/> Universal Resource Locators (Web URL's)  |
|  | <input type="checkbox"/> Internet Protocol (IP) addresses   |
|  | <input type="checkbox"/> Biometric Identifiers (includes fingerprints and voiceprints)  |
|  | <input type="checkbox"/> Full face and other identifiable photographic images and any comparable images   |
|  | <input type="checkbox"/> Any other unique identifying number, characteristic or code (may include clinical trial record number, barcode, revealing occupation, unique condition or publicized clinical event) |

I confirm that the information to be used and/or disclosed contains none of these identifiers and that I have no actual knowledge that the information could, alone or in combination, be used to identify any individual subject of the information.

---

Print Name

Signature

Date

---

Project Name: \_\_\_\_\_

Department: \_\_\_\_\_

IRB Protocol Number, if applicable: \_\_\_\_\_

Contact Name: \_\_\_\_\_

A Limited Data Set is information from which “facial” identifiers have been removed. Specifically, as it relates to the individual or his or her relative, employers or household members, all of the following identifiers must be removed in order for health information to be a Limited Data Set.

- Names (including initials)
- Postal address information other than precinct, town, city, state and zip code
- Telephone number
- Fax number
- Email address
- Social Security Number
- Medical Record Number
- Health Plan Beneficiary Number
- Account Numbers
- Certificate or license numbers
- Vehicle identification/serial numbers including license plate numbers
- Device identification/serial numbers
- Universal Resource Locators (Web URL's)
- Internet Protocol (IP) addresses
- Biometric Identifiers (includes fingerprints and voiceprints)
- Full face and other identifiable photographic images and any comparable images

I confirm that the information to be used and/or disclosed contains none of the identifiers listed above. I understand that the person(s) who will receive this Limited Data Set must execute a Data Use Agreement before receiving this information.

---

Print Name

Signature

Date

---