De-identified Data Set

The Privacy Rule permits a covered entity or its business associate to release data that have been de-identified without obtaining an Authorization and without further restrictions upon use or disclosure because de-identified data is not PHI and, therefore, not subject to the Privacy Rule.

A covered entity or business associate may de-identify a data set in one of two methods. The first method, (the “Safe Harbor” method) involves the removal all 18 HIPAA identifiers. In the second method the covered entity formally determines that there is no reasonable basis to believe the data can be used to identify an individual.

Under the second method, the “Expert Determination” method, a qualified statistician—using generally accepted statistical and scientific principles and methods—determines that the risk of re-identification of the individual that is the subject of the information is low. The qualified statistician must document the methods and analysis that justify his/her determination.

The two de-identification methods provided in the Privacy Rule are illustrated below.

**Protected Health Information (Identifiable Data)**

Protected health information (PHI) includes all individually identifiable health information relating to the past, present or future health status, provision of health care, or payment for health care of/for an individual that is created or received by a Covered Entity or Business Associate.

**Privacy Office**

privacyoffice@email.arizona.edu

**Consult the University of Arizona Privacy Office**

privacyoffice@email.arizona.edu to determine whether your project requires a DUA.

**Limited Data Set**

A limited data set is a data set that is stripped of certain direct identifiers specified in the Privacy Rule. A limited data set may be disclosed to an outside party without a patient’s authorization only if certain conditions are met. First, the purpose of the disclosure must be for research, public health, or health care operations purposes. Second, the person or entity receiving the information must sign a data use agreement (DUA) with the covered entity or its business associate.

**Limited Data Sets may include only the following identifiers:**

- Dates, such as admission, discharge, service, and date of birth (DOB)
- City, state, and zip code (not street address)
- Any other unique code or identifier that is not listed as a direct identifier.

This means that in order for a data set to be considered a limited data set, all of the following direct identifiers as they relate to the individual or his/her relatives, employers, or household members must be removed:

- Names
- Street addresses (other than town, city, state, and zip code)
- Telephone and fax numbers
- Email addresses
- Social security numbers
- Medical record numbers
- Health plan beneficiary numbers
- Account numbers
- Certificate/driver’s license numbers
- Vehicle identifiers and serial numbers
- Web Universal Resource Locators (URL)
- Internet Protocol (IP) addresses
- Biometric identifiers
- Full face photographic images and any comparable images
- Any other unique identifying number, characteristic, or code.

**De-identified Data Set**

The Privacy Rule permits a covered entity or its business associate to release data that have been de-identified without obtaining an Authorization and without further restrictions upon use or disclosure because de-identified data is not PHI and, therefore, not subject to the Privacy Rule.

A covered entity or business associate may de-identify a data set in one of two methods. The first method, (the “Safe Harbor” method) involves the removal all 18 HIPAA identifiers. In the second method the covered entity formally determines that there is no reasonable basis to believe the data can be used to identify an individual.

Under the second method, the “Expert Determination” method, a qualified statistician—using generally accepted statistical and scientific principles and methods—determines that the risk of re-identification of the individual that is the subject of the information is low. The qualified statistician must document the methods and analysis that justify his/her determination.

The two de-identification methods provided in the Privacy Rule are illustrated below.

**HIPAA Privacy Rule**

De-identification Method

**Figure from OCR guidance available here:**
https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html

**Consult the University of Arizona Privacy Office**

privacyoffice@email.arizona.edu if you have questions about how to de-identify your data set.
For questions or for more information, please contact the UA Privacy Office at:

520.621.1465 or
privacyoffice@email.arizona.edu