The Network
for Public Health Law

Ideas. Experience. Practical answers.

DE-IDENTIFICATION

Reference

HIPAA Privacy Rule's Safe Harbor De-Identification Method

Remove following identifiers of the individual or of relatives, employers, or household members of the individual:

1. Names
2. 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 the ZIP code if, according to the current publicly available data from the Bureau of the Census:
   a. The geographic unit formed by combining all ZIP codes with the same three initial digits contains more than 20,000 people; and
   b. The initial three digits of a ZIP code for all such geographic units containing 20,000 or fewer people is changed to 000
3. All elements of dates (except year) for dates that are directly related to an individual, including birth date, admission date, discharge date, death date, 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
4. Telephone numbers
5. Fax Numbers
6. Email addresses
7. Social Security numbers
8. Medical record numbers

9. Health plan beneficiary numbers
10. Account numbers
11. Certificate/license numbers
12. Vehicle identifiers and serial numbers, including license plate numbers
13. Device identifiers and serial numbers
14. Web Universal Resource Locators
15. Internet Protocol addresses
16. Biometric identifiers, including finger and voice prints
17. Full-face photographic images and any comparable images
18. Any other unique identifying number, characteristic, or code, except as permitted with respect to a re-identification code

And, the covered entity does not have actual knowledge that the information could be used alone or in combination with other information to identify an individual who is a subject of the information. HIPAA Privacy Rule, 45 CFR §164.514(b)(2).

Covered entities may either remove or transform these identifiers, such as by replacing the identifiers with category names, generic data, symbols, random values or pseudonyms.[1] There is no requirement to remove physician or other workforce or vendor names in the dataset. Covered entities must not utilize parts or derivatives of the identifiers, such as the last four numbers of the Social Security number. Even dates associated with test measures or treatment must adhere to the Safe Harbor Method.

The Safe Harbor method requires removal of unique identifying numbers, characteristics or codes. Examples of an “identifying number” include a medical record number and a clinical trial number. An identifying characteristic “may be anything that distinguishes an individual and allows for identification,” such as a patient’s occupation - “President of the University of Richmond.” An example of an identifying code is the barcode embedded into a patient’s medical record. [2]

Finally, whether or not the covered entity has actual knowledge that the remaining information is identifiable turns on whether it has “clear and direct knowledge that the remaining information could be used, either alone or in combination with other information, to identify an individual who is a subject of the information.”[3]

Examples of actual knowledge include:

1. Where there is a revealing occupation, such as “former president of the State University.” If this information were combined with other data, it would be identifiable.
2. Where the data recipient has a family member in the data and would recognize him or her.
3. Where there is a publicized rare clinical event.
4. Where the organization knows that the anticipated data recipient has a table or algorithm that can identify patients.

A covered entities’ mere knowledge of re-identification studies or attacks is not enough to equate to actual knowledge that the methods would be used against the data.[4]

Safe Harbor does not require execution of a data use agreement.

SUPPORTERS

Robert Wood Johnson Foundation

The Network for Public Health Law is a national initiative of the Robert Wood Johnson Foundation.

This document was developed by Sallie Milam, JD, CIPP/US/G, Deputy Director, Network for Public Health Law – Mid-States Region Office, and reviewed by Denise Chrysler, JD, Director, Network for Public Health Law – Mid-States Region Office. The Network for Public Health Law provides information and technical assistance on issues related to public health. The legal information and assistance provided in this document does not constitute legal advice or legal representation. For legal advice, please consult specific legal counsel.

[1] NIST, De-Identification of Personal Information. NISTIR 8053. Simson L. Garfinkel pp. 15-16. (October 2015).

[2] Guidance on De-identification of Protected Health Information, p. 26. (November 26, 2012). https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/De-identification/hhs_deid_guidance.pdf .

[3] *Id*. at 27.

[4] *Id*. at 27-28.

February, 2019