

HIPPA Deidentification

De-identified data are data that contain none of the 18 HIPAA identifiers listed in the "What is PHI?" section. If all of the 18 identifiers are removed, the information is no longer (1) individually identifiable, (2) PHI, and (3) subject to HIPAA's requirements. A de-identified data set may be coded with a unique identifier that cannot be traced back to the individual for the purpose of being re-identified by the recipient at a later date. De-identified data may include gender, age, race, or relevant information regarding disease or tissue source and can later be re-identified, by the original holder of the data, if necessary, by means of a unique, non identifiable, code for purposes of carrying out research. It is important to remember that re-identification will subject the information to HIPAA's requirements. A researcher must resubmit the protocol to the IRB for approval when re-identification of the data is desired.

A data set may also be considered de-identified if an expert in statistical and scientific methods determines and documents that the methods used to de-identify or code the data present a very small risk that the information can be used alone or in combination with other reasonably available information to identify an individual.

"Anonymous" data are not necessarily considered de-identified under HIPAA. Anonymity under the federal Common Rule requires that individuals cannot be readily ascertained by the investigator and cannot be associated with the data. According to the Common Rule standard, anonymous data may retain dates of treatment. Under HIPAA's more stringent requirements, however, such data would be considered identifiable data.

What is PHI?

PHI is considered individually identifiable if it includes one or more of the following identifiers:

1. Names
2. All geographic subdivisions smaller than a State, including:
 - street address
 - city
 - county
 - precinct
 - zip codes 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.
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. All elements of dates (except year) for dates related to an individual, including:

- birth date
- admission date
- discharge date
- date of death
- 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

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 (URLs)

15. Internet Protocol (IP) address numbers

16. Biometric identifiers, including finger and voice prints

17. Full face photographic images and any comparable images

18. Any other unique identifying numbers, characteristics, or codes