



GUIDANCE DOCUMENT: List of 18 Identifiers and Definition of PHI (Protected Health Information)

List of 18 Identifiers

1. Names;
2. All geographical 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.
3. 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;
4. Phone numbers;
5. Fax numbers;
6. Electronic mail 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 (URLs);
15. Internet Protocol (IP) address numbers;
16. Biometric identifiers, including finger and voice prints;
17. Full face photographic images and any comparable images; and
18. Any other unique identifying number, characteristic, or code (note this does not mean the unique code assigned by the investigator to code the data)

There are also additional standards and criteria to protect individual's privacy from re-identification. Any code used to replace the identifiers in datasets 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, the researcher must not have actual knowledge that the research subject could be re-identified from the remaining identifiers in the PHI used in the research study. 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.



Definition:

What is PHI?

Protected health information (PHI) is any information in the medical record or designated record set that can be used to identify an individual and that was created, used, or disclosed in the course of providing a health care service such as diagnosis or treatment. HIPAA regulations allow researchers to access and use PHI when necessary to conduct research. However, HIPAA only affects research that uses, creates, or discloses PHI that will be entered in to the medical record or will be used for healthcare services, such as treatment, payment or operations.

For example, PHI is used in research studies involving review of existing medical records for research information, such as retrospective chart review. Also, studies that create new medical information because a health care service is being performed as part of research, such as diagnosing a health condition or a new drug or device for treating a health condition, create PHI that will be entered into the medical record. For example, sponsored clinical trails that submit data to the U.S. Food and Drug Administration involve PHI and are therefore subject to HIPAA regulations.

What is not PHI?

In contrast, some research studies use data that is person-identifiable because it includes personal identifiers such as name, address, but it is not considered to be PHI because the data are not associated with or derived from a healthcare service event (treatment, payment, operations, medical records) not entered into the medical records, nor will the subject/patient be informed of the results. Research health information that is kept only in the researcher's records is not subject to HIPAA but is regulated by other human subjects' protection regulations.

Examples of research health information not subject to HIPAA include such studies as the use of aggregate data, diagnostic tests that do not go into the medical record because they are part of a basic research study and the results will not be disclosed to the subject, and testing done without the PHI identifiers. Some genetic basic research can fall into this category such as the search for potential genetic markers, promoter control elements, and other exploratory genetic research. In contrast, genetic testing for a known disease that is considered to be part of diagnosis, treatment and health care would be considered to use PHI and therefore subject to HIPAA regulations.

Also note, health information by itself without the 18 identifiers is not considered to be PHI. For example, a dataset of vital signs by themselves do not constitute protected health information. However, if the vital signs dataset includes medical record numbers, then the entire dataset must be protected since it contains an identifier. PHI is anything that can be used to identify an individual such as private information, facial images, fingerprints, and voiceprints. These can be associated with medical records, biological specimens, biometrics, data sets, as well as direct identifiers of the research subjects in clinical trials.