HIPAA

What does it take to de-identify data?

De-identifying data requires the removal of the 18 HIPAA identifiers. The standard requires removal of all direct (e.g., name, MRN, SSN) and indirect (e.g., ZIP code and dates related to health) identifiers. In particular, for research, we cannot include dates or study identification numbers in a de-identified dataset. Dates can be listed as a year, but this may not be useful information for research. Durations ("time until") can be listed, but zip codes can only be listed as the first three digits. Please send any questions about de-identifying data to Lawrence "Doc" Muhlbaier (lawrence.muhlbaier@duke.edu or 919-668-8774).

What are the 18 HIPAA identifiers

1. Names.
2. All geographic subdivisions smaller than a state, including street address, city, county, precinct, ZIP Code, and their equivalent geographical codes, except for the initial three digits of a ZIP Code if, according to the current publicly available data from the Bureau of the Census:
   - The geographic unit formed by combining all ZIP Codes with the same three initial digits contains more than 20,000 people.
   - The initial three digits of a ZIP Code for all such geographic units containing 20,000 or fewer people are 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. Telephone numbers.
5. Facsimile numbers.
6. Electronic mail addresses.
7. Social security numbers.
8. Medical record numbers.
9. Health plan beneficiary numbers.
10. Account numbers.
12. Vehicle identifiers and serial numbers, including license plate numbers.
15. Internet protocol (IP) address numbers.
16. Biometric identifiers, including fingerprints and voiceprints.
17. Full-face photographic images and any comparable images.
18. Any other unique identifying number, characteristic, or code, unless otherwise permitted
by the Privacy Rule for re-identification.