# De-Identified Data Sets and Limited Data Sets

The following chart describes the information that must be *eliminated* from a database, registry, or any other data set for the data set to be considered "de-identified" or a "limited data set". Appropriately de-identified data sets are not regulated by HIPAA. Limited data sets may be used or disclosed for research, public health, and other limited purposes, but only by those who sign a "data use agreement" (available from the Privacy Director, IRBMED, or DRDA). Note that for each data element listed below, the information must be eliminated with respect to the patient *and* to any of the patient's relatives, employers, or household members.

*Important: Even if HIPAA does not regulate the use of a dataset or permits its use or disclosure for research, federal regulations and University policies governing human subjects research may still apply. Contact IRBMED for more information at irbmed@umich.edu.*

| Data Element | De-Identified Data Set[*] | Limited Data Set |
|---|---|---|
| Names | Remove | Remove |
| Address, city and other geographic information smaller than state. *3-digit zip code may be included in a de-identified data set for an area where more than 20,000 people live; use "000" if fewer than 20,000 people live there.* | Remove | Remove postal address information other than city, town, state or zip code. |
| All elements of dates (except year); plus age and any date (including year) if age is over 89. *Examples: date of birth, date of death, date of admission, date of discharge, date of service.* | Remove | May be included. |
| Telephone, fax numbers; e-mail addresses, web URL addresses, IP addresses. | Remove | Remove |
| Social security number, medical record number, health plan beneficiary number, any account number, certificate or license number. | Remove | Remove |
| Vehicle identifiers and serial numbers, including license plate numbers. | Remove | Remove |
| Device identifiers and serial numbers. | Remove | Remove |
| Biometric identifiers (e.g., fingerprints; voice prints). *DNA is not considered a biometric identifier for purposes of HIPAA.* | Remove | Remove |
| Full-face photographs and any comparable images. | Remove | Remove |
| Any other unique identifying number, characteristic or code. | Remove[†] | May be included. |

---

[*] Even if all of the information listed in this column is removed, if the researcher knows that any remaining information in the data set could be used to re-identify a patient (e.g., a diagnosis code where the disease is very rare), then the data set is not considered de-identified.

[†] If links must be maintained in the data set for potential later re-identification, they must be completely unrelated to any of the above elements. For example, a patient's initials or a scrambled social security number are not permitted in a de-identified data set. A subject code that reflects the order in which subjects were enrolled into a trial would be permitted.