HIPAA 'Protected Health Information': What Does PHI Include? 

HIPAA.com has received from its readers requests for information on topics related to HIPAA Administrative Simplification Privacy and Security Rules and to updates to those rules reflected in the HITECH Act provisions of the American Recovery and Reinvestment Act of 2009, signed by President Obama on February 17, 2009. Of particular interest to readers is: what exactly is protected health information (PHI)?

Protected Health Information

To get to protected health information, you have to examine two definitions that were in Section 1171 of Part C of Subtitle F of Public Law 104-191 (August 21, 1996): Health Insurance Portability and Accountability Act of 1996: Administrative Simplification. These statutory definitions are of health information and individually identifiable health information.

"Health information means any information, whether oral or recorded in any form or medium, that–

(A) is created or received by a health care provider, health plan, public health authority, employer, life insurer, school or university, or health care clearinghouse; and

(B) relates to the past, present, or future physical or mental health or condition of any individual, the provision of health care to an individual, or the past, present, or future payment for the provision of health care to an individual."

"Individually identifiable health information is information that is a subset of health information, including demographic information collected from an individual, and:

(1) Is created or received by a health care provider, health plan, employer, or health care clearinghouse; and

(2) Relates to the past, present, or future physical or mental health or condition of an individual; the provision of health care to an individual; or the past, present, or future payment for the provision of health care to an individual; and

(i.) That identifies the individual; or

(ii.) With respect to which there is a reasonable basis to believe the information can be used to identify the individual."

Protected health information is defined in 45 CFR 160.103, where ‘CFR’ means ‘Code of Federal Regulations’, and, as defined, is referenced in Section 13400 of Subtitle D (‘Privacy’) of the HITECH Act.

"Protected health information means individually identifiable health information [defined above]:

(1) Except as provided in paragraph (2) of this definition, that is:

(i.) Transmitted by electronic media;

(ii.) Maintained in electronic media; or

(iii.) Transmitted or maintained in any other form or medium.

(2) Protected health information excludes individually identifiable health information in:

(i.) Education records covered by the Family Educational Rights and Privacy Act, as amended, 20 U.S.C. 1232g;

(ii.) Records described at 20 U.S.C. 1232g(a)(4)(B)(iv); and

(iii.) Employment records held by a covered entity in its role as employer."

The HIPAA Privacy Rule covers protected health information in any medium while the HIPAA Security Rule covers electronic protected health information.

With those definitions in place, the question becomes: what elements comprise protected health information such that if they were removed, items (i) and (ii) of (2) in the definition of individually identifiable health information would not obtain. The answer is in the de-identification standard and its two implementation specifications of the HIPAA Privacy Rule [45 CFR 164.514]:

(a) "Standard: de-identification of protected health information. Health information [defined above] that does not identify an individual and with respect to which there is no reasonable basis to believe that the information can be used to identify an individual is not individually identifiable health information.

(b) Implementation specifications: requirements for de-identification of protected health information. A covered entity may determine that health information is not individually identifiable health information only if:

(1) A person with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable:

(i) Applying such principles and methods, determines that the risk is very small that the information could be used, alone or in combination with other reasonably available information, by an anticipated recipient to identify an individual who is subject of the information; and

(ii) Documents the methods and results of the analysis that justify such determination; or

(2) The following identifiers of the individual or of relatives, employers, or household members of the individual, are removed:

(A) Names;
(B) 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 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.
(C) 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;

(D) Telephone numbers;

(E) Fax numbers;

(F) Electronic mail addresses;

(G) Social security numbers;

(H) Medical record numbers;

(I) Health plan beneficiary numbers;

(J) Account numbers;

(K) Certificate/license numbers;

(L) Vehicle identifiers and serial numbers, including license plate numbers;

(M) Device identifiers and serial numbers;

(N) Web Universal Resource Locators (URLs);

(O) Internet Protocol (IP) address numbers;

(P) Biometric identifiers, including finger and voice prints;

(Q) Full face photographic images and any comparable images; and

(R) Any other unique identifying number, characteristic, or code, except as permitted by paragraph (c) of this section; and

(ii) 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.

(c) Implementation specifications: re-identification. A covered entity may assign a code or other means of record identification to allow information de-identified under this section to be re-identified by the covered entity, provided that:

(1) Derivation. The code or other means of record identification is not derived from or related to information about the individual and is not otherwise capable of being translated so as to identify the individual; and
(2) Security. The covered entity does not use or disclose the code or other means of record identification for any other purpose, and does not disclose the mechanism for re-identification.”

With HHS’s release of the Interim Final Rule, ‘Breach Notification for Unsecured Protected Health Information,’ published in the Federal Register on Monday, August 24, 2009, note the following: "If information is de-identified in accordance with 45 CFR 164.514(b) [the first implementation specification, defined above], it is not protected health information, and thus, any inadvertent or unauthorized use or disclosure of such information will not be considered a breach for purposes of this subpart.” [74 Federal Register 42743]