Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule

This page provides guidance about methods and approaches to achieve de-identification in accordance with the Health Insurance Portability and Accountability Act of 1996 (HIPAA) Privacy Rule. The guidance explains and answers questions regarding the two methods that can be used to satisfy the Privacy Rule's de-identification standard: Expert Determination and Safe Harbor[1]. This guidance is intended to assist covered entities to understand what is de-identification, the general process by which de-identified information is created, and the options available for performing de-identification.

In developing this guidance, the Office for Civil Rights (OCR) solicited input from stakeholders with practical, technical and policy experience in de-identification. OCR convened stakeholders at a workshop consisting of multiple panel sessions held March 8-9, 2010, in Washington, DC. Each panel addressed a specific topic related to the Privacy Rule's de-identification methodologies and policies. The workshop was open to the public and each panel was followed by a question and answer period. Read more on the Workshop on the HIPAA Privacy Rule's De-Identification Standard. Read the Full Guidance - PDF.

General

Guidance on Satisfying the Expert Determination Method

Protected Health Information

The HIPAA Privacy Rule protects most “individually identifiable health information” held or transmitted by a covered entity or its business associate, in any form or medium, whether electronic, on paper, or oral. The Privacy Rule calls this information *protected health information* (PHI)[2]. Protected health information is

information, including demographic information, which relates to:

- the individual's past, present, or future physical or mental health or condition,
- the provision of health care to the individual, or
- the past, present, or future payment for the provision of health care to the individual, and that identifies the individual or for which there is a reasonable basis to believe can be used to identify the individual. Protected health information includes many common identifiers (e.g., name, address, birth date, Social Security Number) when they can be associated with the health information listed above.

For example, a medical record, laboratory report, or hospital bill would be PHI because each document would contain a patient's name and/or other identifying information associated with the health data content.

By contrast, a health plan report that only noted the average age of health plan members was 45 years would not be PHI because that information, although developed by aggregating information from individual plan member records, does not identify any individual plan members and there is no reasonable basis to believe that it could be used to identify an individual.

The relationship with health information is fundamental. Identifying information alone, such as personal names, residential addresses, or phone numbers, would not necessarily be designated as PHI. For instance, if such information was reported as part of a publicly accessible data source, such as a phone book, then this information would not be PHI because it is not related to heath data (see above). If such information was listed with health condition, health care provision or payment data, such as an indication that the individual was treated at a certain clinic, then this information would be PHI.

Back to top

Covered Entities, Business Associates, and PHI

In general, the protections of the Privacy Rule apply to information held by covered entities and their business associates. HIPAA defines a covered entity as 1) a health care provider that conducts certain standard administrative and financial transactions in electronic form; 2) a health care clearinghouse; or 3) a health plan.[3] A business associate is a person or entity (other than a member of the covered entity's workforce) that performs certain functions or activities on behalf of, or provides certain services to, a covered entity that involve the use or disclosure of protected health information. A covered entity may use a business associate to de-identify PHI on its behalf only to the extent such activity is authorized by their business associate agreement.

See the OCR website http://www.hhs.gov/ocr/privacy/ for detailed information about the Privacy Rule and how it protects the privacy of health information.

Back to top

De-identification and its Rationale

The increasing adoption of health information technologies in the United States accelerates their potential to facilitate beneficial studies that combine large, complex data sets from multiple sources. The process of de-identification, by which identifiers are removed from the health information, mitigates privacy risks to individuals and thereby supports the secondary use of data for comparative effectiveness studies, policy assessment, life sciences research, and other endeavors.

The Privacy Rule was designed to protect individually identifiable health information through permitting only certain uses and disclosures of PHI provided by the Rule, or as authorized by the individual subject of the information. However, in recognition of the potential utility of health information even when it is not individually identifiable, §164.502(d) of the Privacy Rule permits a covered entity or its business associate to create information that is not individually identifiable by following the de-identification standard and implementation specifications in §164.514(a)-(b). These provisions allow the entity to use and disclose information that neither identifies nor provides a reasonable basis to identify an individual.[4] As discussed below, the Privacy Rule provides two de-identification methods: 1) a formal determination by a qualified expert; or 2) the removal of specified individual identifiers as well as absence of actual knowledge by the covered entity that the remaining information could be used alone or in combination with other information to identify the individual.

Both methods, even when properly applied, yield de-identified data that retains some risk of identification. Although the risk is very small, it is not zero, and there is a possibility that de-identified data could be linked back to the identity of the patient to which it corresponds.

Regardless of the method by which de-identification is achieved, the Privacy Rule does not restrict the use or disclosure of de-identified health information, as it is no longer considered protected health information.

Back to top

The De-identification Standard

Section 164.514(a) of the HIPAA Privacy Rule provides the standard for de-identification of protected health information. Under this standard, health information is not individually identifiable if it does not identify an individual and if the covered entity has no reasonable basis to believe it can be used to identify an individual.

§ 164.514 Other requirements relating to uses and disclosures of protected health information.
(a) *Standard: de-identification of protected health information.* Health information 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.