HIPAA Compliance for Open Data Policy

Background

The American Heart Association (AHA) open data policy requires that grant applicants include a data sharing plan as part of the application process for most programs. The data sharing plan must safeguard identifying information related to research subjects in order to comply with the Health Insurance Portability and Accountability Act of 1996 (HIPAA) Privacy Rule. The HIPAA Privacy Rule establishes the conditions under which "protected health information" may be used or disclosed. 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. 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 protected health information. Indeed, many of these details are already widely available on the Internet. However, tying this information directly to a health condition or treatment plan, for example, would push this into the realm of protected health information.

Compliance via De-Identification

The United States Department of Health and Human Services has developed guidance for methods to achieve de-identification in compliance with the HIPAA Privacy Rule. These methods fall into two tracks,) 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. It is the second of these paths (deemed the "Safe Harbor" method by HHS) that is most relevant to the Heart Association's nascent open data policy.

The Safe Harbor method of de-identification requires the removal of 18 specific elements from the research data:

1. Names
2. 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 the ZIP code if,

according to the current publicly available data from the Bureau of the Census: (a) The geographic unit formed by combining all ZIP codes with the same three initial digits contains more than 20,000 people; and (b) The initial three digits of a ZIP code for all such geographic units containing 20,000 or fewer people is changed to 000 3.

3. All elements of dates (except year) for dates that are directly related to an individual, including birth date, admission date, discharge date, death date, 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. 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
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
18. Any other unique identifying number, characteristic, code, except as allowed under the ID specifications (§164.514c)

The researcher responsible for the data should also warrant that he/she does not have actual knowledge that the remaining information could be used alone or in combination with other information to identify an individual who is a subject of the information.

The Department of Health and Human Services provides an FAQ that delves into some of the intricacies associated with the Safe Harbor method.

Demonstrating Compliance

As required by the American Heart Association open data policy, the grant application for most programs will include the applicant's proposed data sharing plan. Within this section, the Heart Association includes language that warrants the applicant's compliance with HIPAA regulations.

As data sharing plans come to fruition, Association staff will spot check the datasets to confirm that they adhere to the Safe Harbor guidelines. In the event that HIPAA noncompliance is identified, the datasets in question must be taken down immediately. Grant recipients will then be contacted and referred to the Safe Harbor method details for correction and reposting of their data.