Northwestern University IRB Guidance on Evaluating Reports of Data Incidents

The federal regulations (45 CFR 46.111(a)(7) and 21 CFR 56.111(a)(7)) require that research involving human participants have adequate provisions in place to protect the privacy of participants and maintain data in a confidential manner in order for the research to receive IRB approval. Additionally, the respect for persons and beneficence principles of the Belmont Report requires the IRB to ensure the participants' privacy and confidentiality protected throughout the course of the study.

This purpose of this guidance document is to distinguish between the different types of data, the difference between privacy and confidentiality, and provide guidance on preparing Reportable New Information (RNI) submissions involving data.

Key Definitions

Health Insurance Portability and Accountability Act (HIPAA):

An act passed by Congress in 1996, HIPAA does the following:

- Provides the ability to transfer and continue health insurance coverage for millions of American workers and their families when they change or lose their jobs;
- Reduces health care fraud and abuse;
- Mandates industry-wide standards for health care information on electronic billing and other processes; and
- Requires the protection and confidential handling of protected health information

Personal Health Information (PHI):

Individually identifiable information relating to the past, present, or future health status of an individual that is created, collected, or transmitted, or maintained by a HIPAA-covered entity in relation to the provision of healthcare, payment for healthcare services, or use in healthcare operations. Under HIPAA, there are 18 unique identifiers that are considered PHI:

1. Names (Full or last name and initial)
2. All geographical identifiers smaller than a state, except for the initial three digits of a zip code if, according to the current publicly available data from the U.S. 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; and the initial three digits of a zip code for all such geographic units containing 20,000 or fewer people is changed to 000
3. Dates (other than year) directly related to an individual
4. Phone Numbers
5. Fax numbers
6. Email addresses
7. Social Security numbers
8. Medical record numbers
9. Health insurance beneficiary numbers
10. Account numbers
11. Certificate/license numbers
12. Vehicle identifiers (including serial numbers and license plate numbers)
13. Device identifiers and serial numbers;

14. Web Uniform Resource Locators (URLs)
15. Internet Protocol (IP) address numbers
16. Biometric identifiers, including finger, retinal and voice prints
17. Full face photographic images and any comparable images
18. Any other unique identifying number, characteristic, or code except the unique code assigned by the investigator to code the data

Personally Identifiable Information (PII):

Data that could potentially be used to identify a particular person. Examples of PII include a full name, Social Security number, driver's license number, bank account number, passport number, and email address.

What is the difference between PHI and PII?

Information or data from the medical records that can be used to identify an individual. The data are considered to be PHI before an informed consent form is signed. Data are considered to be PII after a consent form has been signed.

See additional information here.

Anonymous:

Data are anonymous if no one, not even the researcher, can connect the data to the individual who provided it. No identifying information is collected from the individual, including direct identifiers such as name, address, etc.

De-Identified:

Data are considered de-identified when any direct or indirect identifiers or codes linking the data to the individual subject's identity are stripped and destroyed. De-identified health information neither identifies nor provides a reasonable basis to identify an individual.

Coded:

Identifying information (such as name) that would enable the investigator to readily ascertain the identity of the individual to whom the private information or specimens pertain has been replaced with a code (number, letter, symbol, or any combination) and a key to decipher the code exists, enabling linkage of the identifying information to the private information or specimens.

Confidential:

Confidential data has a link between the data and the individual who provided it. The research team is obligated to protect the data from disclosure outside the research according to the terms of the research protocol and the informed consent document.

Identifiable:

> ***Direct Identifiers:*** information that can be directly linked to an individual, such as name, address, etc**.**
>
> ***Indirect Identifiers:*** information regarding other unique individual characteristics

Sensitive information:

PHI including psychotherapy notes or information relating to relating to HIV/AIDS; behavioral or mental health; developmental disabilities; treatment for substance (alcohol and/or drugs) use disorder; genetic testing and counseling; artificial insemination; sexual assault/abuse; domestic abuse of an adult with a disability; child abuse and neglect; and, if the individual is a minor, sexually transmitted illnesses, pregnancy and birth control.

Privacy:

A person's desire to control the access of others to themselves. For example, persons may not want to be seen entering a place that might stigmatize them, such as a pregnancy counseling center that is clearly identified as such by signs on the front of the building. Privacy concerns people, whereas confidentiality concerns data.

Confidentiality:

An extension of privacy that refers to the researcher's agreement with the participant about how the participant's identifiable private information will be handled, managed, and disseminated.

What is the difference between privacy and confidentiality?

Privacy ***"People"***	**Confidentiality** ***"Data"***
• The way potential participants are identified and contacted • The setting that potential participants will interact with the researcher team and who is present during research procedures • The methods used to collect information about participants • The type of information being collected • Access to the minimum amount of information necessary to conduct the research	• An extension of privacy • Pertains to identifiable data • An agreement about maintenance and who has access to identifiable data • What procedures will be put in place to ensure that only authorized individuals will have access to the information, and • Limitations (if any) to these confidentiality procedures • In regards to HIPAA, protection of patients from inappropriate disclosures of Protected Health Information (PHI)

Data Incident:

An event that leads to a violation of data policies and puts sensitive data at risk of exposure. This is a broad term that includes many different kinds of events.

Data Breach:

A breach is, generally, an impermissible use or disclosure under the Privacy Rule that compromises the security or privacy of the protected health information. An impermissible use or disclosure of protected health information is presumed to be a breach unless the covered entity or business associate, as applicable, demonstrates that there is a low probability that the protected health information has been compromised based on a risk assessment of at least the following factors:

- The nature and extent of the protected health information involved, including the types of identifiers and the likelihood of re-identification;
- The unauthorized person who used the protected health information or to whom the disclosure was made;
- Whether the protected health information was actually acquired or viewed; and
- The extent to which the risk to the protected health information has been mitigated.

Preparing RNI Submissions that Involve Data Incidents

When reporting an event that involves data privacy, confidentiality, and/or a data incident, it is important to distinguish between research data and clinical data, and whether or not HIPAA is involved. Prior to reporting an event, appropriate contacts should be notified of the event and consulted as needed.

- **Carl Cammarata, CFE**
  Senior Director, Chief Information Security Officer
  Feinberg School of Medicine
  carl.cammarata@northwestern.edu
- **Abby Cosentino-Boehm, DrPH**
  Director of Clinical Research Operations
  Feinberg School of Medicine
  a-cosentino-boehm@northwestern.edu

If the event involves Shirley Ryan AbilityLab (SRALAB), the following contact should also be notified:

- **Melissa Mitchell, JD, CHC, CPC**
  Corporate Compliance Officer
  Shirley Ryan AbilityLab
  mmitchell2@sralab.org

Please ensure that the submission is as detailed as possible – provide as much information as possible.

Events should be reporting in real-time, as soon as possible, even if a solution is pending.

Guidance on how to prepare a corrective and preventive action (CAPA) plan can be found here**.**

Additional guidance on data retention and storage from the Office of Research and NUIT can be found here:

- Research Data Ownership, Retention, and Access Policy
- Guidelines for Security and Confidentiality of Data Files
- Data Access Policy
- HIPAA/ISO Information Security Guidance
- Protect Your Research Environment Guidance

Resources:

https://aspe.hhs.gov/report/health-insurance-portability-and-accountability-act-1996

https://www.hhs.gov/hipaa/for-professionals/index.html

https://www.hhs.gov/hipaa/for-professionals/special-topics/hitech-act-enforcement-interim-final-rule/index.html

https://www.research.uci.edu/compliance/human-research-protections/researchers/privacy-and-confidentiality.html

https://www.it.northwestern.edu/index.html

https://www.research.northwestern.edu/policies/