What is and is not Protected Health Information (PHI) in Research Settings:

Research-related Health Information (RHI) and its relation to HIPAA PHI

A position paper of the University of Calfornia Systemwide HIPAA Implementation Taskforce

The Privacy Rule of the Health Insurance Portability and Accountability Act of 1996 (HIPAA) creates a set of requirements and restrictions for the handling of Protected Health Information (PHI). PHI is defined as a subset of individually identifiable health information (IIHI) that is maintained or transmitted in any form, including oral communications that is created or received by a health care provider, relates to the past, present or future physical or mental condition of an individual; provision of health care to an individual; or payment for that health care; and identifies or could be used to identify the individual. HIPAA specifically recognizes that PHI may be created, used and disclosed in the course of performing research.

The broad definition of IIHI has led some to conclude that any individually-identifiable fact about a person arising out of their participation in a research study would be PHI if it had immediate or potential relevance to normal or abnormal functioning (ie., health and disease) at a molecular, physiologic, or functional level. However, life sciences research is by its nature a wide-ranging discovery activity where the importance of individual observations, and the incremental development of new experimental methods, is pursued to contribute to generalizable knowledge. In many if not most cases it is simply not known whether the research results will be significant, correct, and relevant to healthcare services or to the health and well being of a particular individual. A large fraction of the biomedical research involving human subjects that is sponsored by NIH and other federal and not-for-profit entities is done to characterize and better understand disease processes without an associated intervention designed to correct them.

Therefore, we propose a definition for Research-related Health Information (RHI) which shares some characteristics of HIPAA PHI, but would be governed by a different set of principles and best practices that respect the rights of individuals while at the same time catalyzing progress in biomedical and behavioral sciences. The key distinction between RHI and PHI is that PHI is associated with or derived from a healthcare service event. Thus, research studies that use medical records as a source of person-identifiable research data are using PHI, and interventional clinical studies where treatments are being compared for safety and effectiveness would create PHI. In contrast, a research study that does not include a diagnostic or therapeutic intervention, and does not acquire health-related facts about a person by copying them from a medical record, would create information that if individually identifiable would be considered RHI.

An example of this would be a study of brain imaging in schizophrenia designed to correlate imaging patterns with participant symptoms, where appropriately-consented participants might provide facts about their medical history by interview or filling out research data forms. Since this data was provided as part of voluntary participation in a study, and not as a byproduct of a healthcare service event, it would be governed by the

principles of respect for persons enumerated in the federal Common Rule (45 CFR 46), including the maintenance of confidentiality and security of the information. Ethical standards for the conduct of research would apply, including the concepts of need to know, and minimum necessary. Best practices for RHI also include the separation of person-identifiable demographic data from scientific data associated with a unique study identifier, and the provision of additional protections such as encryption and role-based access control for individually-identifiable data elements in the research record. This approach supports common scientific procedures such as statistical analysis based on study identifier while protecting the confidentiality of individuals. But best practices for RHI would not include HIPAA's administrative requirements for business partner agreements, logging of disclosures, audit trails and right to request amendment of records.

The table below summarizes the characteristics of research data that would be considered PHI and research data that would be considered RHI.

Characteristic	HIPAA PHI	RHI
Individually identifiable ie., meets HIPAA definition of IIHI	Yes	Yes
Used for support clinical decision making for an individual, or for payment or operations	Yes	No
Associated with healthcare service event	Yes	No
Need-to-know, minimum necessary access control	Yes	Yes
Separation of person-identifiable and non-person identifiable data elements wherever feasible	No	Yes
Individual authorization (consent) for creation and use of data	Varies by use	Yes
Business Partner agreements for disclosures	Yes	No
Logs and audit trails of use and disclosure	Yes	Consistent with current best practice for research records
Right to request amendment of records	Yes	At discretion of investigator

The concept of Research-related Health Information recognizes that the HIPAA Privacy Rule was crafted for records associated with an individual's health care, and that in some instances health care records may be used or produced in the course of doing research.

RHI defines a related but distinct class of information arising from biomedical and behavioral research that is not associated with health care service provided to an individual, for which there are similar principles of confidentiality but fewer administrative and documentation requirements. When RHI and PHI are admixed in a research project, it may become impossible to determine the source and use of a particular item of information or data, thus it would seem prudent to apply PHI standards to any project where even a fraction of the research records are derived from or include PHI.