Research Use of Protected Health Information (PHI) Extracted from Epic
May 18, 2020

Is there a difference between Protected Health Information (PHI) and Individually Identified Health Information (IIHI)?
By definition, PHI is IIHI that is subject to HIPAA. IIHI is referred to as PHI when it is created, received, or stored by a HIPAA covered entity. PHI is subject to HIPAA whereas IIHI is not. Depending on the data, there may be other regulatory and ethical requirements for the use and sharing of IIHI. Note that IIHI becomes PHI when held by a HIPAA covered entity and PHI becomes IIHI outside of a HIPAA covered entity. The regulatory requirements are based on who is holding the data rather than where the data originates and the regulatory obligations do not flow with the data once it leaves the covered entity unless there is a contractual obligation to do so such as a Business Associate Agreement (see below).

When must a researcher in a department not subject to HIPAA, nonetheless comply with the regulation?
HIPAA regulations are applicable only in some areas of the University (for a complete list see https://hipaa.yale.edu/sites/default/files/files/List%20of%20Covered%20Departments%202019.pdf). Researchers in departments not subject to HIPAA are not mandated by law or University policy to comply with HIPAA policies. However, in some cases, researchers may be collaborating with a colleague who is subject to HIPAA and may agree to abide by HIPAA Privacy and Security policies or promise to secure the data as PHI in an IRB protocol. In such cases, these researchers would be expected to ensure that PHI is only used or shared as allowed under the approved IRB protocol and that any devices that will create, access, receive or store PHI meet the minimum security standards for high risk data (see https://cybersecurity.yale.edu/minimumsecuritystandards).

Are there requirements for researchers who receive health information in a department not subject to HIPAA and who have not agreed to abide by HIPAA?
While HIPAA would not be applicable outside the HIPAA covered departments, IIHI may nonetheless be high risk data depending on the type of data and associated identifiers. Yale's data classification policy should be consulted to determine the risk classification of the data which will then determine the appropriate security controls needed for devices that will access, store, transmit, or receive the data (https://cybersecurity.yale.edu/protectyourdata). Generally speaking, only the most benign health information would be considered medium or low risk when associated with patient identifiers.

When is health information considered to be identified?
Different regulations have different definitions of when data is considered to be identified. Under HIPAA, data is identified if it is associated with any of the 18 defined identifiers (see https://hipaa.yale.edu/sites/default/files/files/5039-EX-De-ID-EXA.pdf). Data is considered to be deidentified if either 1) all 18 identifiers are removed; or 2) a qualified statistician certifies

that the standard for statistical de-identification is meet (see https://hipaa.yale.edu/sites/default/files/files/5039-PR.pdf). Note that HIPAA considers any of the identifiers alone, when derived from a clinical source, to be PHI subject to HIPAA when held by a HIPAA-covered department. This would include for example:

- a list of patient names without any diagnosis
- a patient's medical condition along with a date of service for clinical care
- a patient's diagnosis along with their zip code.

The IRB regulations (the Common Rule), does not define "deidentified" but rather refers to identifiable private information which is private information for which the identity of the subject is or may readily be ascertained by the investigator or associated with the information. The disconnect between HIPAA and the Common Rule means that research data may be considered identifiable under HIPAA but not under the Common Rule.

To further complicate the matter, data subject to the EU's General Data Protection Regulation (GDPR), California Consumer Privacy Act (CCPA), Family Educational Rights and Privacy Act (FERPA), etc. all have different standards. If your project is subject to additional privacy constraints, please contact the Privacy Office at privacy@yale.edu

Are there constraints on how de-identified data can be used or shared?
De-identified data is not constrained by HIPAA or the Common Rule and may be used for research by University researchers without an IRB approval. De-identified data may not be shared with non-Yale researchers or organizations without an appropriate use agreement in place. The Office of Sponsored Projects can assist in ensuring an appropriate agreement is in place.

What is a Limited Data Set?
A Limited Data Set (LDS) is PHI that has most identifiers removed and is subject to a HIPAA Data Use Agreement (HIPAA-DUA). LDS's are allowed to keep most dates and some geographic information (see https://hipaa.yale.edu/sites/default/files/files/5039-EX-LDS-EXB.pdf). LDS's may be used for research purposes without patient authorization when approved by an IRB and with a signed HIPAA-DUA. Note that when the LDS is derived from University or Health System data and will not be shared with individuals unaffiliated with the University, an internal HIPAA-DUA is available at https://hipaa.yale.edu/sites/default/files/files/5039-FR-Internal-Data-Use-Agreement.pdf and should be stored with the study records for at least 6 years.

How do I know what data I'm approved to access?
The IRB protocol should provide a clear and detailed description of the data to be extracted from the medical record. The request must meet the Minimum Necessary standard which means that only the minimum data needed for the research will be collected. When requesting a data report from the Joint Data Analytics Team (JDAT), the IRB protocol will be reviewed and only that data described in the IRB protocol may be provided.

My study involves a data set that was pull from Epic under an IRB approval with a waiver of patient authorization. May I use the patient identifiers to contact patients?
You may only use data pulled from the medical record to contact patients when explicitly approved by the IRB. The Yale IRB rarely approves patient contact by someone who is not involved in the patient's care. The IRB protocol would need to have clearly indicated that medical record information would be used to contact patients and the IRB would need to state in the approval letter that the contact was approved.

My study was approved as a medical record review protocol. May I use the patient data collected to contact patients?
No. Medical record reviews are presumed to be limited to review and to not involve patient contact. If patient contact is needed, a standard IRB protocol should be submitted. The IRB protocol would need to have clearly indicated that medical record information would be used to contact patients and the IRB would need to state in the approval letter that the contact was approved.

I have access to Epic and am performing a review of records preparatory to research. May I record patient contact information to contact the patients for recruitment into my study?
Patient contact for the purpose of research may only be done under an approved IRB protocol with explicit approval from the IRB to contact the patients. Reviews preparatory to research are limited to determining if the study is feasible or preparing a grant or IRB protocol. Data accessed under a review preparatory to research may not be recorded in an identified form and may not be used for recruitment. Recruitment procedures require IRB approval and either patient authorization or a partial IRB waiver of authorization for recruitment.

When is a Business Associate Agreement (BAA) needed to share data? For example, is it needed to share data with departments not covered by HIPAA such as YSPH?
BAAs are used when PHI will be shared with an individual or company that is not part of Yale and who will be performing activities on our behalf. The terms of the agreement require that the Business Associate comply with HIPAA and limit how they can use the data. BAAs are generally not appropriate for research collaborations but may be appropriate for venders working on behalf of the study who will have access to or store PHI such as cloud service providers, transcriptionists, subject recruitment venders, contractors, etc.

My study was granted an exemption from the IRB. Do I still have HIPAA obligations?
A determination that a study is exempt from IRB oversight does not change the HIPAA status of any underlying data. If the study will involve PHI, HIPAA requirements related to the sharing and securing the data will still apply. Note that because the definitions in the Common Rule and HIPAA are different, study data that is not "recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained directly or through identifiers linked to the subjects" may still be considered identifiable under HIPAA and subject to University HIPAA Privacy and Security policies.

Am I required to use the Joint Data Analytics Team (JDAT) to obtain a list of patients with the appropriate diagnosis for recruitment or analysis in my study?
Yale University and Yale New Haven Health allow patients to opt out of research use of their health information. Reports generated by JDAT will automatically exclude any patients who have opted out of research. Researchers who choose not to use JDAT are expected to confirm prospective patient's opt out status in the patient demographic fields in Epic.

Is the approval of the Yale University IRB, including a waiver of HIPAA authorization, sufficient for access to Epic data or do I also need YNHHS approval?
The University IRB serves as the IRB for both the University and Yale New Haven Hospital. A waiver of authorization approved by the University IRB is valid for use of PHI of Yale School of Medicine and Yale New Haven Hospital. Access to data from other YNHHS facilities may require additional review. Note that PHI may only be used as approved in the IRB protocol. Any subsequent use or sharing of the data may require review by the IRB or Privacy Office.

May I share my data set with other researchers or companies once I am done or if I remove identifiers?
Subsequent use or sharing of PHI requires review by an IRB or Privacy Office to ensure any regulatory or ethical obligations are met.

Where can I get more information?

- Yale's HIPAA website: https://hipaa.yale.edu/
- Yale Human Research Protection Program website: https://your.yale.edu/research-support/human-research