Dartmouth Health SYNERGY Informatics Research Data Handbook

Version 1.5 (10-28-2023)

Contents

Introduction and Acknowledgments

The first version of the SYNERGY Informatics Research Data Handbook was posted to the SYNERGY website in November 2023. The purpose of the handbook is to standardize and organize Dartmouth Health knowledge about the acquisition and use of data for research purposes.

Versions

We recognize that the Handbook is incomplete, and we will offer newer versions at least bi-annually. The latest version will always be posted to the Dartmouth SYNERGY website: https://synergy.dartmouth.edu

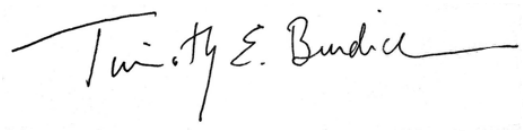
Acknowledgments

The Handbook was the result of many people meeting weekly over several months to reach consensus on the outline and the content:

- Office of Research Operations – IRB: Tonya LaClair, Michael Hill, Cristina Ferrazzano Yaussy, Shannon Sewards
- Office of Research Operations – Informatics: Jen Snide, Robyn Mosher
- Clinical Informatics: John Mecchella, Todd Morrell
- Information Services – Analytics Institute: Carole Felone
- Privacy Office: Rebecca Malila

Of special note, Caren Saunders was invaluable as Project Manager (Office of Research Operations) in keeping the team moving forward.

With gratitude for their dedication to DH research,



Tim Burdick MD MSc MBA

Associate Chief Research Officer for Informatics, Dartmouth Health

Associate Professor and Vice Chair of Research, Community and Family Medicine

Protected Health Information (PHI)

Protected Health Information (PHI) has a very specific definition. It is extremely important to understand PHI since mistakes can lead to significant Federal fines. PHI is any information in the medical record or designated record set that can be used to identify an individual and that was created, used, or disclosed in the course of providing a healthcare service such as diagnosis or treatment.

An identified data set is any collection of data with any one or more of these 18 identifiers:

1. Names;
2. All geographical subdivisions smaller than a State, including street address, city, county, precinct, zip code, and their equivalent geocodes, except for the initial three digits of a zip code, if according to the current publicly available data from the Bureau of the Census: (1) The geographic unit formed by combining all zip codes with the same three initial digits contains more than 20,000 people; and (2) The initial three digits of a zip code for all such geographic units containing 20,000 or fewer people is changed to 000.
3. All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; 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. Phone 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, including license plate 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; and
18. Any other unique identifying number, characteristic, or code (note this does not mean the unique code assigned by the investigator to code the data)

If you are working with data that includes ANY PHI, be sure to follow guidelines relating to Data Use Agreements, Accounting of Disclosures, and proper storage of data in other sections of this handbook

Limited Data Sets vs. De-identified Data Sets

In many cases, it is possible to adjust the data in order to create a limited data set (LDS) or a fully de-identified data set by removing or coding data such as MRN or dates while maintaining the utility of the data.

What is a de-identified data set?

A de-identified data set is different than an LDS. Fully de-identified data sets have *all* 18 PHI elements removed or coded. Sharing a de-identified dataset outside of DH requires a Data Use Agreement.

What is a limited data set (LDS)?

An LDS is one that is only partially de-identified. By definition, an LDS contains some PHI. It is allowed to contain dates (admissions, office visits, lab results, medication orders, etc.); city, state, five-digit zip code; and ages in years, months or days or hours. All other PHI must be removed. Sharing an LDS outside of DH still requires a Data Use Agreement.

Why create a limited or de-identified data set?

There are many, compelling reasons to use a limited or de-identified data set:

1. To protect patient privacy. The study team should conduct the research with the minimum necessary access to PHI.
2. To avoid tracking disclosures. With a limited or fully de-identified data set, you can share the data with other researchers at Dartmouth College and elsewhere without having to track the name of every patient in the dataset under accounting of disclosures.
3. To simplify data use agreements (DUA). Sharing data outside DH requires a signed DUA between DH and the other organization, involving the legal department, delaying the project. DUAs for de-identified datasets are much simpler than for data containing any of the 18 PHI elements.
4. To make storing data easier. The rules regarding where you can store data are much simpler for de-identified data.
5. To reduce the risk and harms of a data breach.

How do researchers obtain a limited or de-identified data set?

Although the real values for all 18 PHI identifiers must be removed, it is possible to create codes for many PHI values. This de-identification can be done by an Analytics Institute Honest Data Broker (HDB) before the study team receives the data.

- **MRN** can be changed to a "fake" or coded numerical identifier. The HDB keeps the "key" which cross-lists the actual MRN with the coded identifier. The research team receives the data with only the coded identifier. [If the study team has a compelling need to know the real patient later (patient safety, statistical analysis, etc.), the HDB can use the key to break the code – usually with additional IRB approval.] If the key is also deleted, and the data can never again be re-identified, the dataset is "anonymized".

Dates can be shifted. For example, if a patient was admitted to the hospital on 7/15/2021, then readmitted on 8/2/2021, the HDB can shift the dates forward by 6 days. The research team would receive the data as "Admitted 7/21/2023" and "Readmitted 8/8/2023". As long as the amount of shift (# days) and the direction of the shift (forward or back) is consistent for all dates for a given patient, the timeline for their care is often still useful for the research project. It is also possible to remove the dates and use a day 0 for an initial (incident) date and then the count of days to all other events. The HDBs

have tools to shift patient dates by a random amount with minimal amount of work. If the study team ever needs the real dates, the HDB can provide that information with IRB modifications.

Do I need to track "accounting of disclosures" for limited or de-identified data sets?

No, you do not need to track which limited or de-identified data set patient records are shared under the accounting of disclosures rule.

Do I need a data use agreement (DUA) for limited or de-identified data sets?

Yes, if you are disclosing data outside of DH, you do need a DUA, even if you have a limited or fully de-identified data set.

Data "Use" vs. "Disclosure" and Accounting of Disclosures (AoD)

Data "use" and data "disclosure" have very different definitions and regulations. It is important that the PI – and all members of their team – understand the distinction and ramifications. When in doubt, please complete the SYNERGY Informatics Consult Intake Form.

What is the "use" of PHI data?

Data "use" covers cases when the data is *not disclosed* outside the DH covered entity. There are two requirements for data to stay within the covered entity:

1. Everybody with access to the PHI data is DH workforce (see definition below); AND
2. The data is stored on a computer inside the DH firewall:
    a. DH server: H: drive or I: drive, or
    b. DH-managed computer local drive (C: drive), or
    c. DH REDCap

Dartmouth College systems are NOT inside the DH firewall, including Granite (which is on a College Server) and College REDCap.

What is a "disclosure" of PHI data?

If the sharing of data does not meet the definition of data "use", then it is considered data "disclosure". A disclosure happens when you share PHI data outside of DH "affiliate covered entity" OR store the data outside the DH firewall. (Limited data sets and de-identified data sets are not considered disclosures. See below.) There is likely a disclosure if either of these is true:

1. One or more persons with access to the PHI data is **NOT** a member of the DH workforce. Students are NOT classified as DH workforce for research purposes despite having a Hitchcock email address, unless specially designated.
2. Data is stored **outside** of the DH IS firewall. Data are inside the DH firewall if stored on a DH H: drive, DH I: drive, DH REDCap (not College REDCap), or a DH-managed computer local C: drive. Anywhere else (Granite server, Dartmouth College computer, etc.) may result in a disclosure.

DH Affiliate Covered Entities and DH Workforce

For the purposes of data use and disclosures, all DH sites and affiliates who are using the Epic EHR (eDH) are part of an "affiliate covered entity" agreement. These locations include DHMC, DH Clinics, the Community Group Practices in southern NH, Alice Peck Day Hospital and clinics, New London Hospital and clinics, Mt. Ascutney Hospital and Health Center, Cheshire Medical Center and clinics, and Visiting Nurse and Hospice of VT & NH. (Southwestern Vermont Medical Center, DH Putnam Physicians, and Valley Regional Hospital (Claremont) will also join the "affiliate covered entity" with Epic (eDH) in the future.) If a researcher is employed by and has an email account from one of these organizations, they are considered "**DH workforce**".

Students are NOT considered to be DH workforce, so their access to data is always considered a "disclosure". This is true for Dartmouth College (including Geisel), Colby-Sawyer, and other academic institutions, even when students are conducting research with a DH team or doing clinical rotations at DH.

What is "accounting of disclosures"?

HIPAA requires that healthcare organizations track every disclosure of PHI (except in limited data sets). Patients may request a list of data that has been disclosed from their medical record – and the reason for the disclosure. The Analytics Institute Honest Data Brokers keep track of all research disclosures in a computer database, affectionately referred to as "feeding the hippo".

Do I need to account for disclosure of limited data sets (LDS) or de-identified data sets?

No, you do not. However, before you decide that you do not need accounting, carefully review information about the 18 PHI elements, as well as the definitions of LDS and de-identified data.

If you are unclear if your data are limited or de-identified, please ask a research data concierge by completing the SYNERGY Informatics Consult Intake Form.

How do I do the accounting of disclosures (AoD)?

If the Analytics Institute Honest Data Brokers provide you with ALL of your data, they will do the accounting for you. In all other cases (e.g., obtaining any PHI data by running your own reports, doing manual chart reviews, using clinic registries, etc.), the Principal Investigator is responsible for making sure that proper AoD happens:

1. Open the Accounting of Disclosures Excel Template
    a. In the top few rows, add the PI Name, the study name, the IRB study number, and the date of AoD, and a few words describing what PHI is released (e.g., "MRN, DOB, chart review of notes, meds and labs")
    b. In the next rows, record the MRN for every chart that was included in your inclusion/exclusion criteria or opened in your manual chart review process. Please send only the MRNs, one per row.
2. Save the Excel File with the following naming convention: Name of PI_AoD_Year-Month-Date. For example: Burdick_AoD_2023-06-19. If you need to update the list, just change the date on the file name.
3. Use ShareFile to securely email the AoD document to Analytics Institute staff at: AI_ResearchData_Information@hitchcock.org . There is a handy ShareFile plug-in for your DH Microsoft Outlook email that makes it quite easy. You can request the plug-in by submitting a DHSM Help Desk ticket.

Data Use Agreement (DUA)

What is a data use agreement (DUA)?
A DUA is a legal contract between two (or more) organizations. One organization agrees to share (disclose) data, and the other organization agrees to receive the data. The DUA usually includes details about the data elements being shared, dates of the data, methods for sharing (transferring) the data, requirements for storing and destroying data, and responsibilities for any data breaches.

When do I need to use a DUA?
Any data disclosed outside of DH may require a data use agreement (DUA). See "Accounting of Disclosure" for definitions of disclosure.

Do I need a DUA if I am sharing (disclosing) a limited data set or de-identified data set?
Yes, you still need a DUA for any data disclosed outside of DH.

What is the Master Data Use and Transfer Agreement (MDUTA)?
The MDUTA is a specific DUA signed by DH and Dartmouth College in order to facilitate the sharing of research data between the two institutions. If you are sharing data from DH to the College, check with CTO/SPA to see if the MDUTA is sufficient or if you need a custom DUA for your project.

How do I get a DUA written and signed?
When you need a DUA, reach out to Sponsored Projects Administration (SPA) or the Clinical Trials Office (CTO). For all studies that are SPA-Managed (nonclinical trial), please use the Research Administration Intake Form For all Industry-related clinical trials data, please use the CTO Intake Form For questions specific to data sharing and study intake process, please refer directly to the ORO Intranet site. [Note: these links are on the DH Intranet and maybe unavailable to non-DH researchers.]

Conducting Manual Chart Reviews

Who may conduct the manual chart review?
Manual chart review must be done by a member of the study team, identified in eIRB and in the protocol, who already has access to eDH. Researchers from DH or Dartmouth College without eDH access should partner with a clinician who can do the chart review. See "Special considerations with students".

Do I need approval before conducting manual chart review for research?
You MUST have approval from 2 groups *before* beginning manual chart review: 1) Research Data Governance group, and 2) the DH IRB. To get started with the process, you should read this handbook thoroughly and then consult a research data concierge by completing the SYNERGY Informatics Consult Intake Form. The data concierge might also ask the Analytics Institute data brokers for their input.

Your study protocol should specify who will conduct the chart review, how you will identify the right patients, exactly what data will be abstracted from charts, and why the Analytics Institute Honest Data Broker (HDB) cannot *reasonably* supply the data. Seek guidance around sharing data outside of DH. You may also need Privacy Office review.

When is manual chart review allowed?
HIPAA law requires the "minimum necessary" chart review that is *reasonable*. Ideally, an Analytics Institute (AI) Honest Data Broker (HDB) creates a research dataset (preferably fully de-identified if possible) for the study team. This way, the study team does not have to do any manual chart review – and the process is faster, more convenient, and in keeping with the HIPAA "minimum necessary" rule. However, there are times when it is unreasonable for the AI HDB to pull data from the data warehouse. Often, the data request would require an extraordinary about of time from the HDB when the research team could get the data themselves more easily. The study team can request permission from the IRB to conduct a limited manual chart review. The IRB staff may reach out to AI HDB for input before making a decision.

Manual chart review WITHOUT Analytics Institute involvement IS ALLOWED if your project meets all 3 criteria:

1) **No PHI disclosed outside of DH:** (a) all members of the study team are DH or affiliate workforce. (Students are NOT DH workforce); ***AND*** (b) data are stored inside the DH firewall: on DH I: or H: drive, DH REDCap, or a DH-managed computer's local C: drive.
2) **The study team can easily identify the right patients** without having to look in lots of random patient charts that are not going to be eligible. For example, the study team may already have a patient registry that meets the inclusion/exclusion criteria reasonably well.
3) **Some data elements cannot reasonably be pulled by an AI HDB**, such as information in clinical notes, imaging reports, or test result comments. Please review the list of data elements easily available to AI HDBs. The IRB staff can make a decision about what is "reasonable" to pull with input from AI HDBs or ORO Informatics as needed.

In this scenario, since no PHI is disclosed outside of DH, you do NOT need to do the accounting of disclosures.

Chart review WITH Analytics Institute involvement

If your project does not meet all 3 criteria above AND you think that manual chart review is necessary, please complete the SYNERGY Informatics Consult Intake Form.

Specific circumstances where consultation or assistance from AI is very helpful to the study team:

1. If PHI data will be disclosed outside of DH, you will need to work with (a) the AI HDBs for accounting of disclosures and (b) work with ORO CTO/SPA on a data use agreement
2. If you don't already have a good list of patients in the cohort, AI HDB may be able to use your inclusion/exclusion criteria to generate a list of patients with MRNs.
3. If some of the research data is easily available in the data warehouse, the HDB can provide a list of patients and a partial dataset. The study team can supplement the dataset with manual chart review.

Waiver of HIPAA Authorization

What is a Waiver of HIPAA Authorization and why is it required? HIPAA requires that patients authorize having their protected health information (PHI) shared with researchers. The IRB can issue a Waiver so that researchers do not need to get patient consent for data sharing. A Waiver is often used for retrospective or secondary data use studies when obtaining consent would be impossible.

How do I get help with a Waiver of HIPAA Authorization? Please consult the Waiver worksheet: HRP-330 - WORKSHEET - HIPAA Authorization.docx (live.com). If you need assistance understanding or requesting the HIPAA Waiver of Authorization, please email IRB@hitchcock.org.

Do I need a Waiver of Authorization for a limited data set or a fully de-identified data? If you are using a Limited Data Set and have a Data Use Agreement, you do NOT need to obtain a Waiver. If you have a fully de-identified data set, you also do not need to obtain a Waiver.

Do I need a Waiver of Authorization if the data is not being disclosed outside of DH? Yes, even if the data is being used only by researchers who are DH workforce (see definition under "DH Affiliate Covered Entities and DH Workforce") *and* the data is stored inside the DH firewall (DH I: drive or H: drive, DH REDCAP, or DH-managed computer hard drive), you still need a Waiver. You also need a Waiver if the data is disclosed outside of DH.

Do I need a Waiver of Authorization for data about Decedents? No, you do not need a Waiver, but the use of decedent data is still subject to other data use rules regarding IRB review, accounting of disclosures, and data use agreements.

Do I need a Waiver of Authorization to use TriNetX? Use of TriNetX does not require IRB determination, approval, or a waiver of authorization because it is fully de-identified.

How do I request a Waiver of HIPAA Authorization from DH IRB? The study team must request a Waiver within the protocol or site supplement.
Include a description of the PHI data elements for which use or access is included in the protocol summary and is necessary for the research.
The protocol must include a detailed explanation for all 3 requirements:

1. How the use or disclosure of protected health information involves no more than a minimal risk to the privacy of individuals, based on, at least, the presence of the following elements:
    a. An adequate plan to protect the PHI identifiers from improper use and disclosure.
    b. An adequate plan to destroy the PHI identifiers at the earliest opportunity consistent with conduct of the research, unless there is a health or research justification for retaining the identifiers or such retention is otherwise required by law.
    c. Adequate written assurances that the PHI will not be reused or disclosed to any other person or entity, except as required by law, for authorized oversight of the research study, or for other research for which the use or disclosure of protected health information for which an authorization or opportunity to agree or object is not required by 45 CFR 164.512.
2. Why the research could NOT practicably be conducted without the waiver or alteration.
3. Why the research could NOT practicably be conducted without access to and use of the protected health information.

Patient Consent to Collection and Disclosure of PHI

In some cases, the research protocol may have patients consent to specific data management. Patient consent MIGHT reduce the obligations under accounting of disclosures, and need for a data use agreement. The written consent document must specify exactly what data will be collected, how it will be collected, with whom the data will be shared, what the researchers will do with the data, and the elements of the HIPAA Authorization.

If your study team is considering using patient consent for a specific use or disclosure of data, please discuss with the DH IRB. The IRB consent template includes the proper language.

Readily Reportable Data for Research

Data elements

This list represents some of the data that is reasonably available from eDH (Epic EHR) from the Analytics Institute. Note that data points captured in eDH change regularly, and may be captured in different parts of the medical record. For items not on this list (text notes, images, scanned media documents, external results and immunizations, etc.) please complete the SYNERGY Informatics Consult Intake Form.

1) Patient demographics
   a) MRN
   b) Name
   c) Sex (assigned a birth and gender identity)
   d) DOB
   e) Race
   f) Ethnicity
   g) Preferred language
   h) Address
   i) Hospital Service Location
   j) RUCA category (rural/urban location scale)
   k) Home phone
   l) Cell phone
   m) Current primary insurance

2) Visit/Encounter data
   a) Encounter date
   b) Age at encounter
   c) Encounter department
   d) Encounter provider and specialty
   e) Visit type
   f) Telehealth category and provider attestation responses

3) Hospital visit data (Inpatient, Outpatient and Emergency)
   a) Admission date
   b) Discharge date
   c) Length of stay
   d) Admitting department
   e) Discharge department
   f) Patient class

4) Flowsheets
   a) Systolic and diastolic blood pressure, heart rate, height, weight, BMI
   b) Pain score

5) Social history
   a) Smoking use (if completed: start date, quit date, # packs per day, # years)
   b) Smokeless tobacco use
   c) Drug use (status and substance)
   d) Birth control use and partner gender identity

6) Conditions
   a) Diagnosis dates
   b) Encounter/Visit diagnosis (ICD—9/10 code and name)
   c) Diagnosis billed in Professional or Hospital Billing
   d) Primary admit or discharge diagnosis for hospital visit
   e) Problem list and Medical History diagnosis (ICD 9/10, name, and date added to chart)
   f) DRG codes (inpatient only)

7) Procedures
   a) Patient age at procedure
   b) Procedure dates and time stamps
   c) Procedure provider
   d) Procedure location (if done at DHMC)
   e) CPT Procedure code/description (ICD-9/10 codes for inpatients only)
   f) Operating room case/log data
   g) Anesthesia data
   h) Implant data

8) Episodes of Care
   a) Pregnancy Episodes data
   b) Transitional Care Management Episodes data
   c) Others

9) EDH Registries
   a) Current membership (y/n)
   b) Earliest activation date
   c) Deactivation date
   d) Appointment/Order Question responses
   e) Question ID/name
   f) Patient response

10) Lab results
   a) Component ID/Name
   b) LOINC (Logical Observation Identifiers Names and Codes) – lab test identifier code
   c) Order date
   d) Result date
   e) Result if numeric/short text
   f) Ordering provider/authorizing provider

11) Imaging data
    a) Image code/name
    b) Order date
    c) Interpretation date
    d) Ordering provider
    e) Order-specific question responses
    f) Interpreter name

12) Medications
    a) Prescribed medication name (brand, generic), dose, instructions, # dispensed, # refills
    b) Medication administration
    c) Medication on reviewed list
    d) Start, stop, refill, and dose adjustment dates
    e) Pharmaceutical class and subclass; Therapeutic class
    f) RxNorm code
    g) Beacon/therapy plans

13) Smart data elements
    a) Smartdata ID/name
    b) Smartdata value (from note, order template, etc.)
    c) Questionnaires
    d) Questions asked
    e) Questions answered
    f) Answers

14) Referrals
    a) Referral order name
    b) Referred to: (internal/external, provider name and address)
    c) Referral order questions
    d) Referring provider name
    e) Referral date and time

Dates data became available in Epic by affiliate organization

As new organizations joined DH, they began contributing Epic data to the Enterprise Data Warehouse for Research:

Organization	**Date data begin in Epic (eDH)**
Lebanon, Manchester, Nashua clinical data	04/2011
Lebanon, Manchester, Nashua billing data	10/2015
Cheshire Medical Center and clinics	11/2017
Alice Peck Day Hospital and clinics	05/2019
New London Hospital and clinics	05/2022

When Do I Need to Have Privacy Office Approval?

You only need to review research data requests with the Privacy Office (PrivacyOffice@hitchcock.org) if it falls into one of these categories:

Substance Use Disorder (SUD) treatment is covered under federal and/or state laws that limit use or disclosure. Data pertaining to care in an SUD treatment clinic, such as the DH Rivermill Addiction Treatment Program, covered under this rule includes dates or locations of treatment, diagnoses, medications and other orders, urine drug testing and other laboratory results, and clinical notes. The rules do NOT cover treatment of SUD outside addiction treatment clinics, such as treatment for SUD in primary care, the Emergency Department, or non-psychiatric med-surg hospital admissions.

Business confidential data may be prohibited from use or disclosure for research purposes. For example, DH has contractual agreements not to share certain financial information such as insurance company reimbursement data or medical device pricing. While these financial or contractual data typically are not a matter of human subjects protection, and even if there is no PHI, DH may be legally obligated to keep such data confidential. There are some situations where this data may be shared for internal use as part of QI/QA initiatives but may not be publicly shared. For projects attempting to show cost reduction with a change in practice, consider utilization rates such as level-of-service or other CPT billing codes as a proxy for cost. Another option is the use of Medicare charge/reimbursement data rather than actual costs/reimbursement data. Examples include (but are not limited to):

- Claim payment data (reimbursement data)
- Certain product costs (ex. Contracted cost of surgical implant devices, etc.)
- Certain compliance data (CMS quality data, The Joint Commission results, etc.)
- Financial data (margin on services or equipment, etc.)

Disclosure of data outside DH or Dartmouth College If you will be sharing data outside of DH, whether it is with Dartmouth College or another institution (e.g., with Vanderbilt), you will need a data use agreement. Sharing with institutions other than Dartmouth College also requires approval of the Privacy Office.

Social Security Numbers are typically not allowed to be used or disclosed for research purposes.

For all other research data requests, you do NOT need routine Privacy Office review.

Research Use of HIV or Genetics Data

NH State Law has strict limitations on the access to HIV or genetics testing data for research purposes. In both cases, researchers are required to obtain written consent from each patient before using any HIV or genetic information. Releasing just the names or identifiers of patients who received HIV or genetic testing must be consented in advance, even if the results of the tests are never divulged.

If your research team is considering accessing or using any HIV or genetic testing data, please allow for lots of time to prepare the study protocol. The study will need to be reviewed by the DH Data Governance Committee before it will be approved by the IRB. Please complete the SYNERGY Informatics Consult Intake Form.

IND/IDE data capture and storage in a 21 CFR 11 compliant system

Who must use a 21 CFR 11 compliant system? All DH Investigators who personally hold an investigational new drug (**IND**) or investigational device exemption (**IDE**) (i.e., regulatory sponsor) must use the DH instance of EDC to capture and store the data collected for their investigational protocols (excludes expanded access IND/IDE protocols). This applies to these studies open to accrual on or after: **October 6, 2023.** Please see the following Policy for DH Institutional requirements regarding the use of EDC: Electronic Data Capture System: Use Requirements for Human Subjects Research (ID: 30016).

What is EDC? The DH instance of Advarra's Electronic Data Capture (EDC) is the system that has been appropriately configured and internally validated to comply with 21 CRF 11. The Office of Research Operations manages the EDC system. The EDC intranet page (currently only available to DH employees) includes a variety of resources for the research community:

- Access requests
- Training requirements
- Workflows and user guides
- Upgrades
- Policies / Procedures

Why did DH implement EDC? *Studies that include an IND or IDE must comply with federal regulation* 21 CFR 11:

- ✓ Includes controls to protect data integrity
- ✓ Enables electronic signatures
- ✓ Supports effective data collection – to streamline analyses
- ✓ Multisite capabilities

What are the key features of EDC? EDC is integrated with eDH (Epic EHR) to support the direct population of specific lab values from eDH into EDC. EDC is integrated with OnCore to reduce duplicate data entry. It is designed to streamline monitoring and validation of data entry and facilitates electronic PI sign-off for eCRFs. EDC also enables DH to serve as a data coordinating center for multisite investigator-initiated trials.

What support will be available to the PI of a study that is required to use EDC? ORO has a dedicated team who will support the designing/building/testing of the database that is built for your protocol. Please see the EDC intranet page (currently available to DH employees only):

Approved Places to Store Research Data

Data used within DH – If you are using data only within DH (not a disclosure of PHI), the preferred data storage is on the DH H: drive, DH I: drive, DH REDCap, or the local hard drive of a DH-managed computer. These storage locations are encrypted and backed up automatically. If the data are fully de-identified (*not a limited data set*), you could store the data on other drives, but these are risky for many reasons. Note that storing data on the Dartmouth College REDCap is considered a disclosure and sharing of data outside of DH.

De-identified data shared to Dartmouth College – If the data is fully de-identified, you may store the data on a DartFS server, Dartmouth College Google drive, or the local hard drive on a personal computer as long as the proper data use agreement (DUA) is in place.

Identified or limited data shared to Dartmouth College – In these cases, the data includes some PHI, so the data must be stored on an approved drive. The preferred location is Granite (see section below on this page for information). Using Granite will facilitate IRB approval and getting a DUA signed. Other locations are possible, with significant caveats. Note that researchers are expressly forbidden from storing any PHI on a DartFS server, including Limited Data Sets.

Data use agreement (DUA) – If you are sharing data with Dartmouth College, regardless of PHI/limited/de-identified, you will need a DUA. The DH/College Master Data Use and Transfer Agreement (MDUTA) might be sufficient, but you should carefully review the DUA section of this handbook. If you aren't sure, check with CTO/SPA as they oversee all DUAs, including the MDUTA.

Storing data not at DH or at Dartmouth College – See section on this topic. If patients provided specific consent for the data sharing AND there is a proper data use agreement in place, the IRB staff may be able to approve the plan. If patients are not providing specific consent for data sharing, you will need to review with the Privacy Office.

What is Granite and who can use it? Granite is a secure, encrypted, HIPAA-compliant computing platform available to Dartmouth Health and Dartmouth College research teams. **It is the preferred place to work on data containing PHI** that is accessed by DH and College researchers and students. Granite is the Dartmouth name for our instance of tiCrypt software. Granite is funded equally by DH and the College. ***Using Granite will make approving your protocol much easier.***

What is DH REDCap and who can use it? *Dartmouth Health* (DH) REDCap is a useful survey and data management platform. *DH* REDCap is approved for data storage, including PHI. There are very few technical limitations preventing PHI being disclosed, so the PI is responsible for making sure people on the study team understand how to use *DH* REDCap responsibly. If you are considering using *DH* REDCap for PHI, please consider Granite first – and complete the SYNERGY Informatics Consult Intake Form for further assistance. [*Dartmouth College* REDCap and Qualtrics are NOT approved for collecting or storing any PHI data without prospective patient consent.]

Table of Places to Store Data

	Data not shared outside DH *(everybody* with data access has a hitchcock.org email)		**Data shared with Dartmouth College (*someone* other than Geisel student with data access does *not* have a hitchcock.org) Data Use Agreement Req'd**		**Notes** To share data outside of DH and Dartmouth College, please contact the DH Privacy Office. Master Data Use & Transfer Agreement might cover.
Data Storage Location	**No PHI** (fully deidentified of all 18 identifiers)	**PHI** (contains 1 or more of 18 identifiers)	**No PHI** (fully deidentified of all 18 identifiers)	**PHI** (contains 1 or more of 18 identifiers)	
DH H: or I: Drive	**Preferred.**	**Preferred.**	Not allowed per access control.	Not allowed per access control.	Secure, encrypted, backed up. College employee and student access to H: or I: only by exception and may require signed Business Associates Agreement
DH REDCap	Allowed.	Allowed with caution. **See Notes.**	DUA required.	Allowed. Use caution.	Since exporting PHI data without safeguards is possible, **consider** consulting SYNERGY Informatics for assistance until more guidance is available. The PI is responsible for managing access and data use in DH REDCap.
DH computer C: drive	Allowed.	Allowed.	NA	NA	Secure, encrypted, but NOT backed up. Only DH computer owner can access data. Not available to non-DH workforce.
EDC (Electronic Data Capture)	NA	**Required for IND/IDE studies.**	NA	NA	EDC is required for data management of IND/IDE studies that require compliance with 22 CFR 11. See the DH EDC intranet page
College Granite	Allowed. Costs may be prohibitive.	Allowed. Costs may be prohibitive.	Not preferred due to cost. DUA required.	**Preferred.** Needs DUA.	***Easiest approval for sharing PHI outside DH*** Maybe too expensive for other use cases. Benefit of access to certain analytics tools even for DH researchers.
College REDCap			DUA required.	**Highly discouraged.** Granite is recommended	Requirements: DH IRB approval, DH Privacy Office approval, Geisel NetID, and DUA. Must discuss risks with College Research Computing in advance.
College Google Drive			DUA required.	**Highly discouraged.** Granite is recommended.	Requirements: DH IRB approval, DH Privacy Office approval, Geisel NetID, and DUA. Must discuss risks with College Research Computing in advance.
College DartFS / Computing Cluster			DUA required.	**Not allowed per policy.**	NOT HIPAA-compliant.

Requesting Data for Quality Improvement vs. Research

Please read "An instrument to differentiate between clinical research and quality improvement" (Ogrinc et al., 2013). See Appendix. The full article is available at the SYNERGY Website.

Are you considering publishing your QI or research project? Although the definition of QI does not depend on plans to publish, you should consider whether or not you plan to publish your project in a peer-reviewed journal. Most journals require that the project be assessed by an IRB for determination of non human subjects research. If you are planning to publish, ***we strongly recommend*** that you submit your project to the IRB for review.

Projects that are clearly QI and will not be published do *not* need to be submitted to the IRB for a determination. For these clear QI projects, it is best practice to create a QI project charter and obtain the sponsorship of your department Chair, Vice President, or other leader.

Projects that are not clearly QI or research should be submitted to the IRB using the HRP-503 NHSR form.

If you think the project is likely QI, you can help the IRB by avoiding research jargon and instead using QI terminology:

Avoid research words	**Use quality improvement (QI) terminology**
Study	QI project
Study Coordinator	QI Project Manager
Participants or Subjects	Patients
Hypothesis	Goal
Intervention	Test of change, PSDA cycle, or Implementation of best practice
Control	Standard care, Current care
Endpoint	Outcome, metric, measure
Knowledge gap	Performance gap

The IRB might make a determination that your project is not human subjects research (NHSR). If the project is NHSR, the IRB will provide a letter stating the project is NHSR, and there will be no further documentation or review required. Alternately, the IRB might decide that the project is HSR, and you will need to follow up with the IRB for project approval.

If you would like to discuss your project before drafting a protocol for the IRB, please complete the SYNERGY Informatics Consult Intake Form.

An instrument to differentiate between clinical research and quality improvement

from Ogrinc G, Nelson WA, Adams SM, O'Hara AE. An instrument to differentiate between clinical research and quality improvement. IRB. 2013 Sep-Oct;35(5):1-8. PMID: 24350502.)

Instrument to Assess the Differences between QI and Clinical Research with Human Subjects

This table is intended to compare and contrast the general characteristics of quality improvement (QI) and clinical research activities and is for use by Institutional Review Boards (IRBs), administrative reviewers, investigators, and improvers. This table is intended to guide discussion among these individuals and is not intended to supplant the judgment of IRBs or local QI ethics review committees. Please start by considering these overarching questions:

1. Will the activities of this project occur within the standard of care? If NO, then proceed to IRB review.
2. Is there risk? If YES, use chart below to determine whether this project requires QI review or IRB review.
3. Is this project primarily intended to generate generalizable knowledge? If YES, proceed to IRB review.
4. Does this project involve vulnerable populations? If YES, use chart below to determine whether this project requires QI review or IRB review.

For each item, choose the column to which the project most closely relates—QI or research. You may only choose one answer. Leave the item blank if neither choice applies.

Attribute	*Quality Improvement*	*Clinical Research with Human Subjects*
Intent and Background	☐ Describes the nature and severity of a specific local performance gap	☐ Identifies a specific deficit in scientific knowledge from the literature
	☐ Focus is to improve a specific aspect of health or health care delivery that is currently NOT consistently and appropriately being implemented at this site	☐ Proposes to address or identify specific hypotheses in order to develop new knowledge or advance existing knowledge
Methods	☐ Mechanisms of the intervention are expected to change over time (i.e., an iterative activity) in response to ongoing feedback	☐ Specific protocol defines the intervention, interaction, and use of collected data and tissues, plus project may rely on the randomization of individuals to enhance confidence in differences
	☐ Plan for intervention and analysis includes an assessment of the system (i.e., process flow diagram, fishbone, etc)	☐ May use qualitative or quantitative methods to make observations, make comparisons between groups, or generate hypotheses
	☐ Statistical methods evaluate system level processes and outcomes over time with statistical process control or other methods	☐ Statistical methods primarily compare differences between groups or correlate observed differences with a known health condition
Intended Benefit	☐ Intervention would be considered within the usual clinician-patient therapeutic relationship	☐ Intervention, interaction, or use of identifiable private information occurs outside of the usual clinician-patient therapeutic relationship
	☐ Direct benefit to participants is indicated (e.g., for the decrease in risk by receiving a vaccination or by creating a safer institutional system)	☐ Direct benefit to each individual participant or for the institution is not typically the intent or is not certain
	☐ Potential local institutional benefit is specified (e.g., increased efficiency or decreased cost)	☐ Potential societal benefit in developing new or advancing existing generalizable knowledge

Attribute	*Quality Improvement*	*Clinical Research with Human Subjects*
Risk	☐ Risk is to privacy or the confidentiality of health information	☐ Risks may be minimal, but may include physical, psychological, emotional, social, or financial risks, as well as risk to privacy or the confidentiality of health information from participation in the project
	☐ Risk may be described as higher for patients by not participating in this activity	☐ The informed consent process describes the risks to participants, who individually and voluntarily decide whether to participate or an IRB grants an alteration or waiver of the consent process
Applicability of Results	☐ Implementation is immediate so that review of results occurs throughout the process and may be used for next QI activity	☐ Results and analysis may be delayed or periodic throughout the duration of the project, except to protect patient safety. The results will primarily be used to inform further investigations, but may be implemented directly into clinical practice.
	☐ Extrapolation of results to other settings is possible, but not the main intent of the activity	☐ Results are intended to generalize beyond the study population
Sharing & Disseminating	☐ System level outcomes, processes, refinement of the intervention, and the applicability of the intervention in specific settings/contexts may be shared through peer-reviewed publication and presentation outside the institution.	☐ It is expected that results will be published or presented to others through a peer-reviewed process

Interpretation

Any checkmarks (even one) in the "Clinical Research" column indicates that there are components of clinical research in the proposed activity. The IRB or local QI ethics review committee should initiate a discussion with the improver/investigator to clarify the proposal. If an activity such as public health practice, program evaluation, or quality improvement includes a research component, then IRB review should occur under current federal guidance and the policies of many institutions.

Explanation and Elaboration of Terms

1. **Vulnerable population.** Any study population that includes students, employees, children, pregnant women, prisoners, active military personnel, individuals who have impaired decision making capacity, or those who are educationally or economically disadvantaged.
2. **Intent.** The state of the investigator's mind that directs the activity.
3. **Quality improvement.** The combined and unceasing efforts of everyone—health care professionals, patients and their families, researchers, administrators, payers, planners, educators—to make changes that will lead to better patient outcome, better system performance, and better professional development.
4. **Clinical research.** A systematic investigation in a clinical setting designed to develop or contribute to generalizable knowledge (The Common Rule definition of research)

Special Considerations with Students

Students are NOT considered to be members of the DH workforce, whether they are from Dartmouth (including Geisel School of Medicine), Colby-Sawyer, or other institutions. This applies to students even if they have a hithcock.org email address, have access to Epic (eDH), are participating on a DH research project, or are doing a clinical rotation at a DH site. Therefore, giving students access to data that includes any PHI is considered a disclosure of data outside of DH. For example, adding a student to a DH REDCap project that includes PHI can change the "use" of data into a "disclosure". This may suddenly trigger new requirements for a data use agreement (DUA), accounting of disclosures, and Waiver of HIPAA Authorization , even though the data is stored inside the DH firewall.