health care systems research network

DUA TOOLKIT

A Guide to Data Use Agreements

Purpose and Description

This guide was created to facilitate the establishment of Data Use Agreements (DUAs) between HCSRN sites. It includes information about:

- When DUAs are needed
- The steps involved in putting a DUA in place
- Tools and resources related to DUAs and PHI disclosures
- Best practices and common pitfalls

TABLE OF CONTENTS

APPENDICES

WHAT IS A DATA USE AGREEMENT?

A Data Use Agreement (DUA) is an agreement that governs the sharing of data between research collaborators who are covered entities under the HIPAA privacy rule. A DUA establishes the ways in which the information in a limited data set may be used by the intended recipient, and how it is protected.

Advantages of a DUA

The HIPAA privacy rule allows a covered entity to use and disclose a limited data set (LDS) for research without obtaining an authorization or a waiver of authorization. A covered entity (e.g., a health plan) may disclose a LDS to another entity or researcher who is not a covered entity when a DUA is in place.

Important upfront considerations

1) Expect that analyses and manuscript authorship will be spread across sites, and ensure all potential authors will have access to data.

DUAs ARE ALWAYS STUDY SPECIFIC

Blanket DUAs do not exist between organizations

Permissions outlined in a DUA

1) Who may receive and use the limited data set
2) Allowable uses and disclosures by the recipient

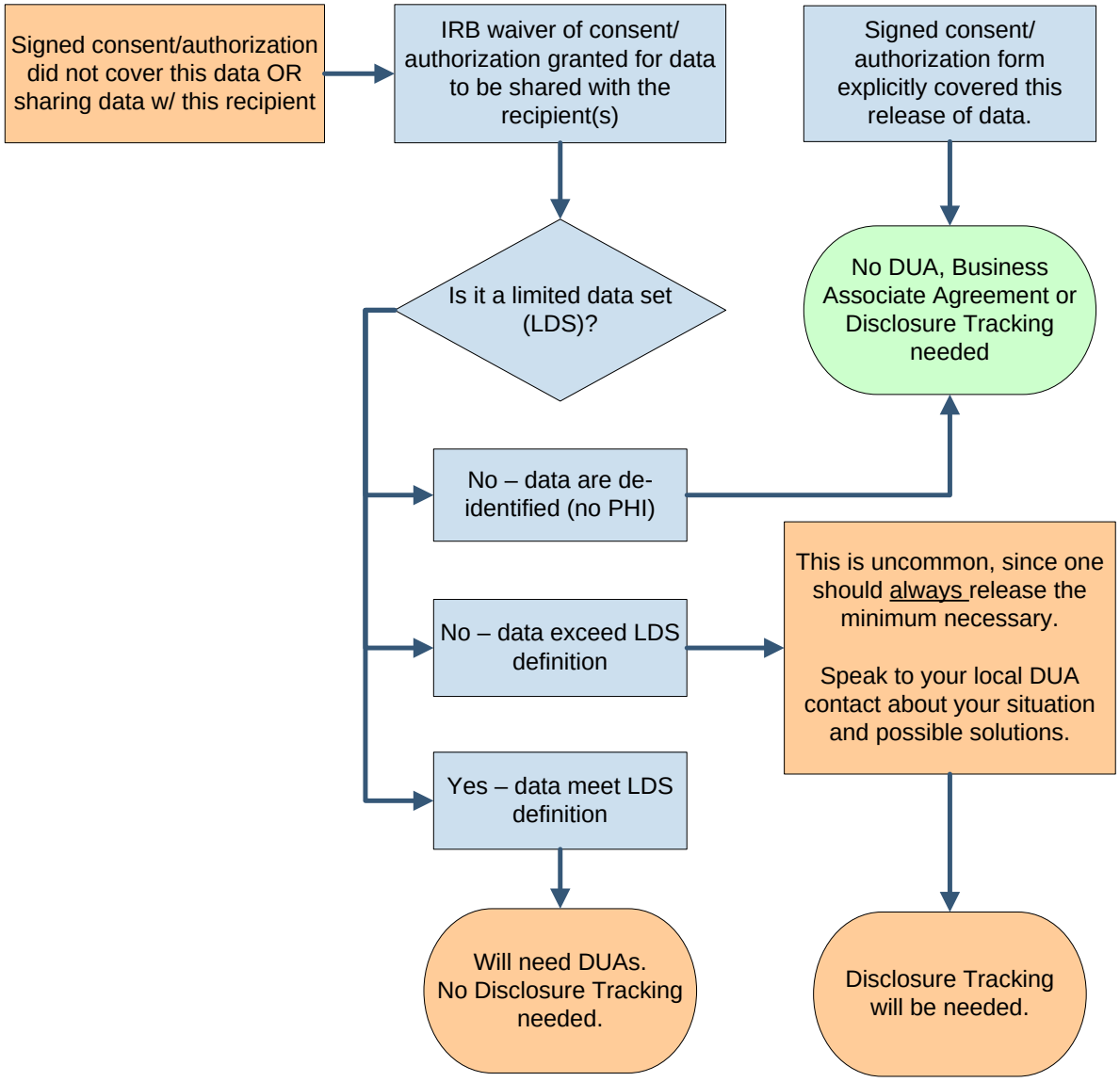
Assurances outlined in a DUA

1) The recipient will not try to identify or contact subjects represented in the LDS.
2) The recipient will not use or disclose/share the data in ways other than stated in the agreement, or as otherwise required by law.
3) The recipient will safeguard the data to prevent such misuse or unauthorized disclosures.
4) The recipient will report any misuse or unauthorized disclosure as soon as known.
5) The recipient will ensure that any agents, including subcontractors, agree and are bound to the restrictions and conditions of the DUA.

WHEN DO I NEED A DUA?

To put it simply, you need a DUA anytime you are sharing data that are not de-identified in a manner that was not explicitly covered in the consent form. Sharing a de-identified data set does not require a DUA, but limited data sets may be shared only after a DUA is in place. The first step is to determine what type of data set you are working with.

Flow Diagram: Do I need a DUA?



Do I have a de-identified data set?

Data are considered de-identified if there's no reasonable way they could be used to identify a person. Thus, de-identified data sets may NOT contain any of the following 18 elements that HIPAA identifies as protected health information (PHI):

1. Names
2. All geographic subdivisions smaller than a state, including street address, city, county, precinct, ZIP Code, and their equivalent geographical codes, except for the initial three digits of a 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
    b. The initial three digits of a ZIP Code for all such geographic units containing 20,000 or fewer people are 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. Telephone numbers
5. Facsimile 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 fingerprints and voiceprints
17. Full-face photographic images and any comparable images
18. Any other unique identifying number, characteristic, or code, unless otherwise permitted by the Privacy Rule for re-identification

Working out the terms of a DUA sometimes takes more time and effort than foreseen.

CONSIDERATION

Is aggregated data or a de-identified data set an option for your study?

Do I have a limited data set?

Limited data sets are NOT de-identified and may contain some (but not all) of the 18 elements that qualify as PHI. For example, a limited data set may NOT include directly identifying information (like name, SSN, or address). However, limited data sets MAY contain the following indirect identifiers:

- town or city, state, zip code;
- ages in years up to 90 years (must aggregate all ages 90 or older);

- dates directly related to an individual – such as birth date, date of death, admission date, discharge date, visit date, diagnosis date, etc. *(Limiting to month/year is preferred).*

A unique study ID can be included in both limited and de-identified data sets – but the number can NOT be an encoded identifier, such as a scrambled birth date, patient initials, last four of social security number, and so on.

My data set exceeds a limited data set--What now?

Remember to release only the minimum necessary data, defined as the least information reasonably necessary to accomplish the intended purpose of the use, disclosure, or request. If you do NOT have a signed written consent authorizing data sharing with the recipient AND you exceed the definition of LDS:

1) Obtain an IRB Waiver of Authorization.

2) Work out contractual solutions between sites, e.g., Business Associate Agreement (BAA), Memorandum of Understanding (MOU), Non-disclosure Agreement, etc.

3) If the data exceed the LDS definition, report both the patients and type of PHI sent outside of the covered entity according to local Disclosure Tracking procedures.

Step 1 and/or 2 may require a great deal of time and resources.

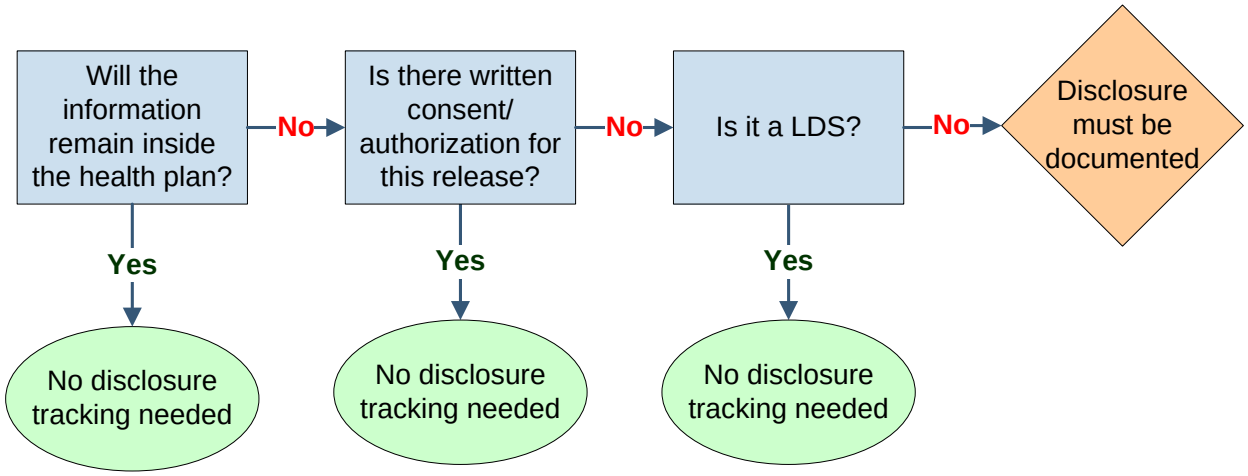
CONSIDERATION

Is it possible to alter your analysis plan so only a LDS is sent?

Disclosure tracking

Disclosures must be tracked any time protected health information is disclosed and either of the following apply:

- Authorization or a waiver of authorization has not been granted.
- Data exceed the definition of a limited data set.



SETTING UP A DUA

This is a general overview of the process of setting up DUAs. There are three overall steps to follow when setting up a DUA:

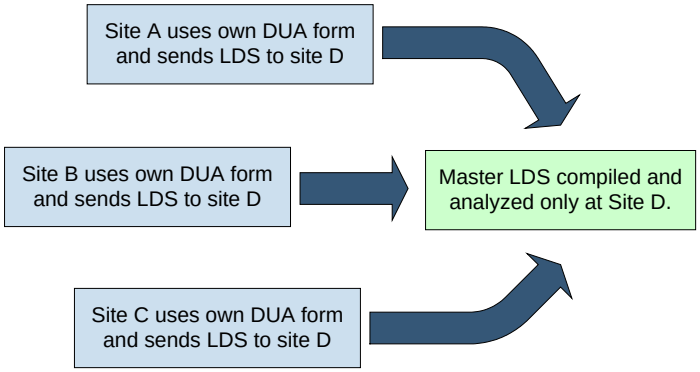
1) Identify the type of DUA needed.
2) Build from a template or previous DUA.
3) Finalize the paperwork.

Step 1: Identify the type of DUA needed

To help illustrate this step, consider the following three common scenarios in multisite studies:

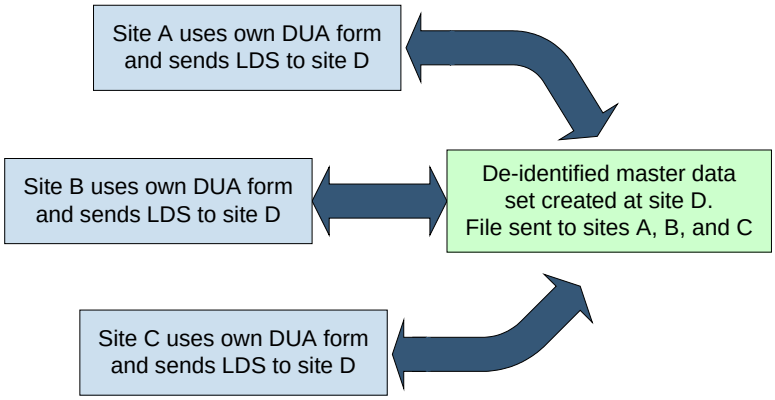
Scenario 1

Sites A, B, C, and D have all collected data in a multi-site study. Site D will create and analyze a master limited data set (LDS), but will NOT send the LDS back to the other sites. Sites A, B, and C need to establish a DUA with site D. Each site will use its own form or an agreed upon template DUA.



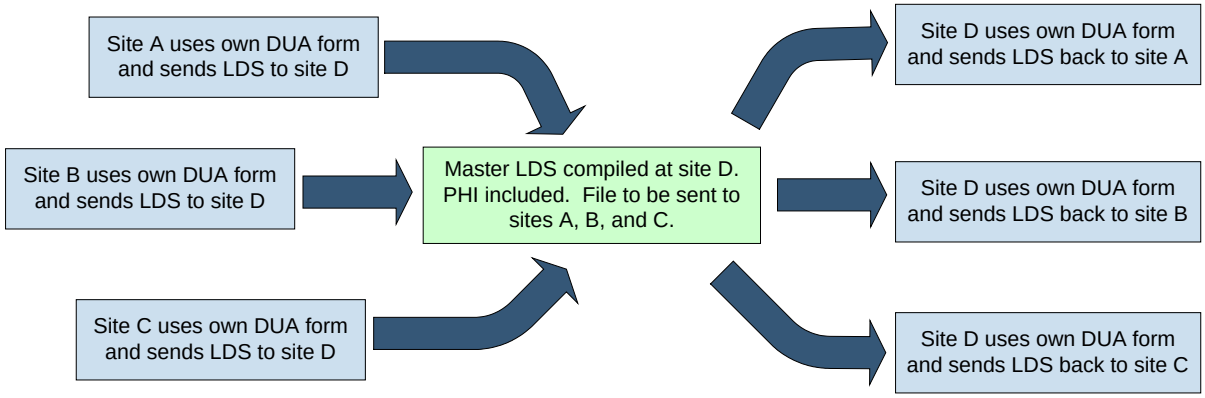
Scenario 2

This scenario is the same as Scenario 1 above, except that site D will compile the LDS and then create a de-identified data set (no PHI) to send back to the other sites for local analyses. Because the data being sent back to sites A, B, and C has been de-identified, no new DUAs are needed.



Scenario 3

This scenario is the same as Scenario 2 above, except that site D will compile the LDS with PHI included to send back to the other sites for local analyses. Site D needs DUAs with A, B, and C before the new LDS can be sent. Site D will use its own DUA form since they house the new master LDS being sent.



Step 2: Build from a template or previous DUA

The HCSRN has developed a pre-negotiated 3-page data use agreement (DUA) template that can be used for projects with (a) straightforward data sharing schemes and (b) that use the HCSRN's pre-negotiated sub-award agreement template.

The HCSRN is also currently pre-negotiating a bidirectional DUA template. Find pre-negotiated HCSRN templates at: http://www.hcsrn.org/en/Tools%20&%20Materials/GrantsContracting/

If you cannot use one of the Network's pre-negotiated templates, it can be helpful to find past or current DUAs between your site and the recipient(s). These may provide useful precedents.

Step 3: Finalize the paperwork

Extra time may be needed for negotiating the terms of the DUA based on the number and nature of the institutions involved, the sensitivity of the data to be shared, the complexity of the data sharing scheme across/between/among sites, and whether or not a pre-negotiated template can be used.

Once the negotiations are completed, the final signed DUA can be distributed to sites for their files.

PROACTIVELY PLANNING FOR SUCCESS

Tips and best practices

1) If data-related issues are already addressed in a subaward agreement, time and resources can be saved when putting together the DUA. The HCSRN's pre-negotiated sub-award agreement template includes many data use and ownership terms up front.

2) Ensure as much time as possible to allow for interpretation and possible reaction to legal wording in the agreements. Set your DUAs up early in the life of the project.

3) Ensure all authors will have access to data. Anticipate opportunities to spread analyses and manuscript authorship broadly across sites and write the DUAs to reflect this.

4) Follow communication pathways set up at individual sites. Circumventing the process causes confusion and adds time.

5) Clarify specific data elements needed for the analysis up front.

6) Required components of a DUA are spelled out in HIPAA. Avoid using a DUA to insert additional requirements that are more appropriate for a contract.

7) Keep the following documents in the project files at each site:

    - Fully signed DUA.
    - Signed Data Release Checklist (or similar documentation required by your site)
    - Documentation of content of the data sent/received (e.g., SAS proc contents report).
    - Cover letter or email documenting data transfer.

Issues commonly leading to delays

Variations in expectations and practices at the local level are a factor in every multi-site study. It can help both Investigators and Project Managers to be aware of the types of problems encountered by others.

- The DUA was written narrowly and uni-directionally. It did not account for the possibility of new analytic plans. For example, only the prime site could send pooled data to subcontractors. The DUA did not address sub-to-sub data sharing for secondary analyses, etc., or the addition of a new site.

- Local interpretation of regulations by legal counsel varied across sites, making mutual agreement much more difficult.
    - Agreement on which state (or site) has jurisdiction, should disputes arise.
    - One site may require more stringent security protections than another site.

- State laws prohibited sites from reaching mutual agreement on some DUA terms.
    - Minnesota, Washington, and Oregon all have state laws pertaining to certain types of data (e.g., the Oregon Genetic Privacy Law) which may supersede federal regulations in the HIPAA Privacy Rule.

- Receipt of aggregated summary data only may preclude certain analyses.

- Sites may hesitate to stray from language used in past DUAs or may not want to make changes to a pre-approved template.

- Trying to involve a non-HCSRN-based Investigator or business associates prolonged negotiations.
    - Example: Data collection or data entry service

- Sites have differing views on the degree of assumed risk to the health plan (e.g., in the event of an unauthorized disclosure) when data are shared.
    - Example: Some health plans may view quality of care data as being a greater risk than data on use of preventive services.

APPENDICES

More about PHI and Data Disclosure

Under HIPAA, the general rule is that researchers must have valid authorization for all uses and disclosures of PHI in connection with research. A valid authorization must include specific elements:

- A description of the PHI being used
- A statement of the purpose of the use of PHI
- A list of those who can use the PHI
- A list of those who can receive the PHI, including the possibility of re-disclosure
- Information about the expiration of the authorization
- Information about the right to revoke the authorization

If an actual expiration date is not provided, then a note pointing this out is required. A statement explaining an expiration event such as the end of the research project is also acceptable.

As to the right to revoke, the authorization must either explain that right or refer to the covered entity's privacy notice, if that is applicable. A revocation must be in writing and can be made at any time, but it may not be effective if a research study has already relied on the authorization. This reliance element only affects information gathered before the revocation and does not allow the entity to disclose PHI after the revocation occurs.

The covered entity – that is, "health plans, health providers and health clearinghouses" or "any entity in the health sector that uses health information in the regular course of business" – may require the authorization as a condition of providing research-related treatment.

If a *limited data set* will be released outside of the covered entity or accessed/used by anyone not employed by the releasing covered entity without a signed authorization or consent form of each individual whose data are used, then documentation of an IRB waiver of authorization must be kept on file by project staff and a DUA signed by the recipient of the data may be required.

If any *PHI beyond a limited data set* will be released outside of the covered entity or accessed/used by anyone not employed by the releasing covered entity without a written authorization signed by each subject whose data are used, then documentation of an IRB waiver of authorization must kept by project staff and project staff must enter pertinent data into a disclosure tracking file. In addition, a business associate agreement may be required.

Is It a covered entity?

A covered entity is a health care provider that conducts certain transactions in electronic form, a health care clearinghouse, or a health plan. A simple way to check if an institution is or is not a covered entity is to look for their HIPAA Notice of Privacy Practices (NOP) on the internet. Covered entities are required to display their NOP.

Is the CDC a covered entity?

No. Although the CDC collects clearinghouse-like data, it is not an agency that handles treatment, payment, and referral transactions for health care providers.

Is the CSS/SEER (the Cancer Surveillance System) a covered entity?

No. CSS/SEER is a Public Health Authority – that is, an agency of the government that is responsible for public health matters as part of its official mandate. The FDA and OSHA are also Public Health Authorities. HIPAA permits covered entities to disclose protected health information, without authorization, to PHAs.

Are covered entities required to document incidental disclosures permitted by the HIPAA Privacy Rule, in an accounting of disclosures provided to an individual?

No. The Privacy Rule includes a specific exception from the accounting standard for incidental disclosures permitted by the Rule. See 45 CFR 164.528(a)(1).

Is It De-identified Data?

May information de-identified under the Privacy Rule's "Safe Harbor" method contain a data element that identifies a time period of less than a year (e.g., the fourth quarter of a specific year)?

No. The Privacy Rule's "Safe Harbor" method for de-identifying health information requires removal of, among other elements, all elements of dates directly related to an individual, except for year. Thus, a data element such as the fourth quarter of a specified year must be removed if a covered entity intends to de-identify data using the "Safe Harbor" method. See section 164.514(b)(1) of the Privacy Rule. From: NIH website

What lab data variables are permitted in a de-identified data set?

De-identified data cannot contain a lab accession number since they can be linked to consumer numbers in health plan data systems. Specimen collection and test dates are not permitted. Considered de-identified are: the year of the date, and the patient's age at the time of the test.

Can a de-identified data set contain an adverse event date?

No. De-identified data sets can contain only the year of the date, not the month or day. However, under HIPAA, there are special considerations for reporting adverse events. If your sponsor is the FDA, you may report adverse events without specific agreements. The minimum necessary standard applies. This would count as a disclosure and would need to be tracked.

Can I send aggregate data without identifiers or dates to a collaborator without putting a DUA in place?

Yes, provided that the likelihood of an individual being re-identified is small. For example if the number in each cell is significant, the data can be shared with other researchers. Even with very small number in a cell, the data may be safe to send, for example if the categories it represents are broad enough, e.g., ages in five- or ten-year groups.

Can I substitute the number of days between a date variable and another date (e.g. randomization date) for the full date of an event to de-identify or limit the data?

Yes, this is one way to de-identify data. But the recipient cannot have the reference date or other information enabling reconstruction of the actual dates. For example, permissible data to send for an immunization study might by Age-in-days-at-MMR#1, Age-in-days-at-MMR#2, and Age-in-days-at-RashDx. This would allow researchers to see if the RashDx occurred within a short time of the vaccinations without ever giving birth date or service dates. If handled in this way, data would be de-identified and could be sent without setting up a DUA. However, if the recipient already has data that would allow them to create dates from the information you send, then you are in fact sending a LDS (even if in piecemeal fashion) and so you would need to set up a DUA.

Is It Disclosure?

An external investigator would like to see paper questionnaires to do some data cleaning. There isn't any personal identifier information on the questionnaires, only a study number. Is there any reason not to send the questionnaires?

Your action will depend upon which variables are on the questionnaire and whether consents and authorizations are in place. Even though the data are on paper, it is still a data set. A DUA could be required. Always check your IRB arrangements before releasing any data that are not de-identified.

If I share provider survey data, is it considered a disclosure?

No, as long as the data do not contain health information. Most provider surveys reflect the beliefs and practices of the provider and are therefore not health information. However, provider surveys may contain sensitive data, so check your IRB arrangements before releasing any data that are not de-identified.

Do we need to account for disclosures of updated contact information on study participants who gave oral consent before April 14, 2003?

It depends. In future studies such disclosures should be tracked. However, in an established study where regular contact with the participant has been maintained, this is not considered a disclosure. An important factor is whether the participants signed a HIPAA authorization form with your health plan - which are sometimes more stringent than HIPAA in categorizing interview results as PHI.

My study includes some subjects who are not health plan enrollees. We have disclosed some PHI on them. Do we need to account for such disclosures?

Yes, you are obligated to account for disclosures of PHI regardless of whether the data pertain to health plan enrollees of your covered entity or other subjects. Your site's disclosure tracking system should have a flag of some kind to mark such disclosures.

GLOSSARY OF TERMS USED

Refer to the Privacy Rule on NIH's website for a complete listing of terms and their specific definitions.

Accounting for Disclosures - Information that describes a covered entity's disclosures of PHI other than for treatment, payment, and health care operations; disclosures made with Authorization; and certain other limited disclosures. For those categories of disclosures that need to be in the accounting, the accounting must include disclosures that have occurred during the 6 years (or a shorter time period at the request of the individual) prior to the date of the request for an accounting. However, PHI disclosures made before the compliance date for a covered entity are not part of the accounting requirement.

Authorization - An individual's written permission to allow a covered entity to use or disclose specified PHI for a particular purpose. Except as otherwise permitted by the Rule, a covered entity may not use or disclose PHI for research purposes without a valid Authorization.

Business Associate - A person or entity who, on behalf of a covered entity, performs or assists in performance of a function or activity involving the use or disclosure of individually identifiable health information, such as data analysis, claims processing or administration, utilization review, and quality assurance reviews, or any other function or activity regulated by the HIPAA Administrative Simplification Rules, including the Privacy Rule. Business associates are also persons or entities performing legal, actuarial, accounting, consulting, data aggregation, management, administrative, accreditation, or financial services to or for a covered entity where performing those services involves disclosure of individually identifiable health information by the covered entity or another business associate of the covered entity to that person or entity. A member of a covered entity's workforce is not one of its business associates. A covered entity may be a business associate of another covered entity.

Covered Entity - A health plan, a health care clearinghouse, or a health care provider who transmits health information in electronic form in connection with a transaction for which HHS has adopted a standard.

Data Use Agreement - An agreement into which the covered entity enters with the intended recipient of a limited data set that establishes the ways in which the information in the limited data set may be used and how it will be protected.

Disclosure - The release, transfer, access to, or divulging of information in any other manner outside the entity holding the information.

Health Care Clearinghouse - A public or private entity, including a billing service, re-pricing company, community health management information system or community health information system, and "value-added" networks and switches that either process or facilitate the processing of health information received from another entity in a nonstandard format or containing nonstandard data content into standard data elements or a standard transaction, or receive a standard transaction from another entity and process or facilitate the processing of health information into a nonstandard format or nonstandard data content for the receiving entity.

Health Care Provider - A provider of services (as defined in section 1861(u) of the Act, 42 U.S.C. 1395x(u)), a provider of medical or health services (as defined in section 1861(s) of the Act, 42

U.S.C. 1395x(s)), and any other person or organization who furnishes, bills, or is paid for health care in the normal course of business.

Health Information - Any information, whether oral or recorded in any form or medium, that (1) is created or received by a health care provider, health plan, public health authority, employer, life insurer, school or university, or health care clearinghouse; and (2) relates to the past, present, or future physical or mental health or condition of an individual; the provision of health care to an individual; or the past, present, or future payment for the provision of health care to an individual.

Health Insurance Portability and Accountability Act of 1996 (HIPAA) - This Act requires, among other things, under the Administrative Simplification subtitle, the adoption of standards, including standards for protecting the privacy of individually identifiable health information.

Health Plan - For the purposes of Title II of HIPAA, an individual or group plan that provides or pays the cost of medical care (as defined in section 2791(a)(2) of the PHS Act, 42 U.S.C. 300gg-91(a)(2)) and including entities and government programs listed in the Rule. Health plan excludes: (1) any policy, plan, or program to the extent that it provides, or pays for the cost of, excepted benefits that are listed in section 2791(c)(1) of the PHS Act, 42 U.S.C. 300gg-91(c)(1); and (2) a government-funded program (unless otherwise included at section 160.103 of HIPAA) whose principal purpose is other than providing, or paying for the cost of, health care or whose principal activity is the direct provision of health care to persons or the making of grants to fund the direct provision of health care to persons.

HHS Protection of Human Subjects Regulations - Regulations intended to protect the rights and welfare of human subjects involved in research conducted or supported by HHS. The HHS regulations include the Federal Policy for the Protection of Human Subjects, effective August 19, 1991, and provide additional protections for pregnant women, fetuses, neonates, prisoners, and children involved in research. The HHS regulations can be found at Title 45 of the **Code of Federal Regulations**, Part 46.

Hybrid Entity - A single legal entity that is a covered entity, performs business activities that include both covered and non-covered functions, and designates its health care components as provided in the Privacy Rule. If a covered entity is a hybrid entity, the Privacy Rule generally applies only to its designated health care components. However, non-health care components of a hybrid entity may be business associates of one or more of its health care components, depending on the nature of their relationship.

Individually Identifiable Health Information - Information that is a subset of health information, including demographic information collected from an individual, and (1) is created or received by a health care provider, health plan, employer, or health care clearinghouse; and (2) relates to the past, present, or future physical or mental health or condition of an individual; the provision of health care to an individual; or the past, present, or future payment for the provision of health care to an individual; and (a) that identifies the individual; or (b) with respect to which there is a reasonable basis to believe the information can be used to identify the individual.

Institutional Review Board (IRB) - An IRB can be used to review and approve a researcher's request to waive or alter the Privacy Rule's requirements for an Authorization. The Privacy Rule does not alter the membership, functions and operations, and review and approval procedures of an IRB regarding the protection of human subjects established by other Federal requirements.

Limited Data Set - Refers to PHI that excludes 16 categories of direct identifiers and may be used or disclosed, for purposes of research, public health, or health care operations, without obtaining either an individual's Authorization or a waiver or an alteration of Authorization for its use and disclosure, with a data use agreement.

Minimum Necessary - The least information reasonably necessary to accomplish the intended purpose of the use, disclosure, or request. Unless an exception applies, this standard applies to a covered entity when using or disclosing PHI or when requesting PHI from another covered entity. A covered entity that is using or disclosing PHI for research without Authorization must make reasonable efforts to limit PHI to the minimum necessary. A covered entity may rely, if reasonable under the circumstances, on documentation of IRB or Privacy Board approval or other appropriate representations and documentation under section 164.512(i) as establishing that the request for protected health information for the research meets the minimum necessary requirements.

Protected Health Information - PHI is individually identifiable health information transmitted by electronic media, maintained in electronic media, or transmitted or maintained in any other form or medium. PHI excludes education records covered by the Family Educational Rights and Privacy Act, as amended, 20 U.S.C. 1232g, records described at 20 U.S.C. 1232g(a)(4)(B)(iv), and employment records held by a covered entity in its role as employer.

Use - With respect to individually identifiable health information, the sharing, employment, application, utilization, examination, or analysis of such information within the entity that maintains such information.

Waiver or Alteration of Authorization - The documentation that the covered entity obtains from a researcher or an IRB or a Privacy Board that states that the IRB or Privacy Board has waived or altered the Privacy Rule's requirement that an individual must authorize a covered entity to use or disclose the individual's PHI for research purposes.

Workforce - Employees, volunteers, trainees, and other persons whose conduct, in the performance of work for a covered entity, is under the direct control of the covered entity, whether or not they are paid by the covered entity.