NIH
Distributed Research Network

POLICIES AND PROCEDURES

10/23/2013

Table of Contents

1. Introduction

1.1 Purpose of this document

This document provides an overview of the NIH Health Care System (HCS) Distributed Research Network (DRN) governance policies, querying capabilities, and responsibilities. The policies and procedures described are developed by the NIH DRN Governance Advisory Committee and approved by the HCS Collaboratory Steering Committee, and others as directed by the HCS Collaboratory Coordinating Center housed at the Duke Clinical Research Institute (DCRI). These policies apply to NIH DRN activities, providing guidance on collaboration in the use of electronic health data, while also safeguarding protected health information and proprietary data. ***These policies do not address or replace procedures and governance of individual research projects funded as part of the NIH HCS Research Collaboratory. Policies regarding funding for NIH DRN activities are addressed elsewhere.***

1.2 Background

The NIH Collaboratory's Electronic Health Records (EHR) Core is responsible for creating and overseeing the NIH DRN. The NIH DRN is described in detail here: https://www.nihcollaboratory.org/Pages/distributed-research-network.aspx.

In brief, the NIH DRN is a collaboration enterprise comprised of software, policies, and procedures that facilitate research partnerships through secure distributed querying of health data held and secured by the NIH DRN Data Partners. The network is administered by the NIH DRN Coordinating Center (CC), which also provides query support and other services to facilitate use of the Network.

The NIH DRN can facilitate the formation of research partnerships by enabling requestors (those requesting information) to identify and contact potential Data Partners (organizations that maintain electronic health data) to identify opportunities for collaboration on specific projects. Data Partners may include those inside or outside the NIH DRN. When requestors and Data Partners agree to collaborate, they work together to answer the questions. The network is best used to facilitate multi-site research and does not preclude initiation of partnerships or conduct of research outside the network, even if research partners are identified using the network.

This document focuses primarily on policies and procedures for querying to support identification of potential research partnerships.

2. Organizational Structure

The EHR Core reports to the HCS Collaboratory Steering Committee. The Core is composed of individuals appointed by the HCS Collaboratory Steering Committee. Overall responsibilities of the EHR Core include developing and operating the NIH DRN, engaging with the Phenotypes/Data Standards Core, identifying ways to assist the Demonstration Projects, and providing guidance to the HCS Collaboratory Steering Committee. The EHR Core also oversees the activities of the NIH DRN Coordinating Center as described below.

3. Roles and Responsibilities

3.1 NIH DRN Coordinating Center

The NIH DRN CC resides in the Department of Population Medicine (DPM) at the Harvard Pilgrim Health Care Institute (HPHCI) and is responsible for the day-to-day operations of the NIH DRN. NIH DRN CC staff members have expertise in the use of electronic health data, epidemiology, health services research, statistics, software development, and operation of distributed networks. The activities of the NIH DRN are overseen by the EHR Core. NIH DRN CC responsibilities include:

- Query Support
    - Develop and oversee the query request, approval, and response process
    - Provide NIH staff and other requesters with support regarding use of the Network, querying capabilities, data resources, and interpretation of results. Support also includes data fitness for use and epidemiologic expertise.
    - Review submitted requests for approval
    - Ensure that all data requests are tested and appropriate for use within the Network and with the specific Data Partners involved (see Query Capacity and Fulfillment)
- Query Capacity and Fulfillment
    - Create the capability and expertise to use publicly available tools, including, but not limited to, those developed by the FDA Mini-Sentinel project (www.minisentinel.org), by ESPnet (www.esphealth.org), and the existing querying features of the open source version of PopMedNet (www.popmednet.org). Tools available include the Mini-Sentinel modular programs, summary table queries, and other programs based on the Mini-Sentinel Common Data Model and ESPnet data model.
    - Capacity to process an agreed upon number of requests, by type of request, per month, as directed by NIH (See Attachment A)
    - Work with Data Partners to establish their engagement and agreements/ contracts and assist with query response
    - The NIH DRN CC will establish a response and reimbursement agreement with Data Partners regarding the process for approving and responding to queries. Agreements will address:
        - Procedures that ensure Data Partners' ability to review and approve queries and results
        - Identification of pre-approved query types (e.g., Mini-Sentinel modular programs)
        - Expectations regarding timeliness of responses
        - Compensation by NIH
        - Review and editorial rights of reports and publications, including but not limited to data provided, shown, and interpretations
        - Appropriate acknowledgement by investigators on publications and reports

      Separate procedures, timelines, and compensation schedules will be developed for: 1) requests that use pre-approved query types and return only aggregate results; and 2) requests that use new programs or require additional work on the part of Data Partners. Examples of pre-approved request types include requests that use Mini-Sentinel's summary tables or modular programs, or ESP menu-

driven queries. Data Partners will have different sets of pre-approved query types; that is, not all pre-approved query types will be available at all Data Partner organizations.

    - Work with Data Partners who have data models for which pre-approved query types do not currently exist (e.g., i2b2, HMORN VDW, OMOP) on: 1) ways to participate in these activities; and 2) develop pre-approved queries that will be acceptable to data partners.

- An NIH DRN CC representative serves as the Network Administrator. The Network Administrator establishes all users on the NIH DRN and assigns each user his/her role, as defined in the NIH DRN governance document (see Section 3.4). The Network Administrator also deletes users, as needed (per communication received by DataMart administrators), and is responsible for the overall functioning of the Network.
- Network Maintenance and Software
    - Incorporate software updates and bug-fixes
    - Secure hosting and maintenance
    - Expand the Network's capabilities (e.g., expanded meta-data capture and search functions, ability to store, tag, and share analytic code with proper access controls)
- Additional Activities
    - Collaborate with the NIH Collaboratory's Data Sharing Work Group to help implement uniform data sharing principles and use of analytic datasets
    - Request information from Data Partners related to data availability and fitness for use
    - Write distributed programs to be run at Data Partner sites, both for quality assurance and approved queries
    - Approve, prioritize, test, and execute approved requests
    - Develop and oversee the query request and response process
    - Ensure that documentation on Data Partners' site characteristics, data availability, and quality are available internally and publicly
    - Ensure all data queries/requests have a protocol/research proposal, appropriate IRB approval, and are submitted by an approved requestor
    - Provide Data Partners with detailed descriptions of each request and the intended use of the results, including, as appropriate, a description of relevant protocols, IRB approval, and approved requestor

3.2 Data Partners

Data Partners are member organizations that have electronic health data resources and the technical capacity to respond to a query via the NIH DRN. Some of the Data Partners have the ability to respond to rapid-response queries. Data Partners are not required to have data stored in a particular format or data model. Data Partners will identify which data models they support (e.g., Mini-Sentinel, i2b2, HMORN VDW, OMOP) and what types of pre-approved requests they are willing and able to process. The NIH DRN CC will only send requests to Data Partners who have self-identified as having the ability and willingness to respond to such requests.

Each NIH DRN Data Partner possesses data acquired through its primary business activities (referred to herein as "original source data"), including but not limited to administrative medical and pharmacy claims data, outpatient and inpatient electronic health records (EHRs),

demographic information, outpatient pharmacy dispensings, and registry data. In addition, some Data Partners possess research data sets and will wish to make those resources available for querying. This may include data developed as part of a clinical research trial or other activity that generates a reusable analytic dataset. Each Data Partner will retain physical and operational control over its data and manage and store the data in accordance with its own institutional policies.

3.2.1 Data Partner Expectations

Expectations for Data Partners include:

- Identify local staff to serve as Site Principal Investigator (PI), Site project manager, and Site analyst
- Maintain local data model(s)
- Meet all of the compliance requirements applicable to their local sites
- Assign one or more local analyst to receive and respond to queries/requests
- Respond to all queries/requests per agreed upon guidelines (see Attachment A)
- Maintain a list of its current staff who are authorized to participate in the NIH DRN, their contact information, and their roles and responsibilities within the Network
- Obtain and maintain local approval related to participation in NIH DRN activities
- Obtain necessary human subjects approvals, as appropriate, based on the request and project (e.g., when recruitment of study participants is required)
- Install and maintain the PopMedNet software
- Notify the Network Administrator at the NIH DRN CC when any authorized user should have his/her access removed (e.g., due to a change in employment or role change)

3.2.2 NIH DRN Data Partner PIs

The NIH DRN Data Partner PIs:

- Meet all agreement/contractual requirements and ensure that applicable local Site compliance requirements are met. This includes, but is not limited to, assuring that secure data storage and transfer are conducted in accordance with local, state, and federal regulations, as well as in accordance with institutional policies and procedures.
- Create a process for deciding whether or not to participate in a query/request
- Respond in a timely and complete manner to NIH DRN CC queries/requests (see Attachment A)
- Communicate clearly and in a timely manner to the NIH DRN CC if they choose to “opt out” of an activity

3.2.3 NIH DRN Data Partner Project Managers

Data Partner project managers are expected to:

- Liaise between the NIH DRN CC and their Sites
- Attend relevant meetings
- Communicate with their local Site PIs to ensure applicable local Site compliance requirements are met

3.2.4 NIH DRN Data Partner Analysts

Data Partner analysts will:

- Respond to queries/requests from the NIH DRN CC within the specified timeframe, as directed by the Site PI or local approval workflow
- Provide documentation to the NIH DRN CC on Site-specific data issues
- Follow established procedures for data queries/requests
- Attend calls and meetings, as needed

3.3 Requestors

Authorized requestors include NIH Collaboratory leadership and the Data Partners, and other individuals/organizations designated by NIH Collaboratory leadership. All requestors must adhere to the responsibilities and expectations outlined below. All queries/requests will be reviewed by the NIH DRN CC. If approved by the NIH DRN CC, all queries/requests will be submitted by the NIH DRN CC staff to the Data Partners through the NIH DRN secure Query Tool.

Requestors are responsible for:

- Completing request forms
- Clearly describing the nature of the request and the intended use of the findings, including grant or other funding applications
- Responding to requests for clarification from the NIH DRN CC and the EHR Core
- Obtaining necessary approvals by IRBs and HIPAA privacy boards
- Working with NIH DRN CC and Data Partners to execute necessary agreements/contracts

Requestors are expected to:

- Use results provided only for the stated and approved purpose. This may include a requirement to keep results confidential.
- Abide by any other limitations on use, issued by the EHR Core, the NIH DRN CC, and/or the Data Partner providing and reporting data

3.4 Network Roles

The following are Network roles that are assigned by the Network Administrator based on Network policies.

1. **Network Administrator:** in accordance with the NIH DRN Governance Advisory Committee, can add approved member organizations; create groups and roles; add/delete Network participants; re-set passwords; and view all queries submitted. The Network Administrator has complete access to the Network and all its functions.

   This role is assigned to the System Host and Software Manager and one or more representatives (or delegates) of the NIH DRN Coordinating Center.

2. **Group Administrator:** is able to review, aggregate, and release results for the group. A group of data partners can designate a person as the group administrator and select rules that require the group administrator to review group results before the results are released to the requestor. Results can be released individually or as an aggregate. Networks do not require a Group Administrator; this role is for convenience of affiliated groups of data partners.

3. **DataMart Administrator**: manages the DataMart preferences on the Portal and local DataMart Client (e.g., what data can be queried and by whom). There can be one or more DataMart Administrators per Data Partner. DataMart Administrators cannot send queries to other DataMarts. A DataMart Administrator manual is available to provide instructions on how to use and interact with the NIH DRN. DataMart Administrators are responsible for communicating with the Network Administrator to make requests for adding or removing participants (e.g., removing access for terminated employees).

   Each Data Partner must assign one or more DataMart Administrators and provide the Network Administrator with the names and contact information for each.

4. **Observer**: users who have rights to view query activity within specific projects, including query descriptions, parameters, and statuses. Observers also have rights to view network information, including participating organizations, users, and DataMarts. Observers do not have the rights to view query results.
5. **Requestor**: can submit queries to DataMarts that have given them or their member organization permission to submit queries (see Section 3.3).

4. Policies and Procedures

4.1 Data Queries/Requests

The procedures outlined below are designed to balance efficiency and Data Partner autonomy. The NIH DRN CC can help determine which kind of query/request is most appropriate for each situation.

For each query/request, there are four steps: initiation, approval, execution, and reporting. Regardless of the approach, the following information will be required from the requestor(s): details for the request, purpose of request, intended use of results, expected level of PHI in the response, and date by which information is needed. All requests must be submitted on the standardized form to the NIH DRN CC, which tracks origination and final disposition of all requests. (See 4.2 Data Request Fulfillment)

4.1.1 Data Completeness and Data Characterization Activities

The NIH DRN Coordinating Center will issue queries/requests to ensure proper system functionality and to help characterize data available within the NIH DRN. These requests will only apply to data that the Data Partners maintain in a supported data model. Results of these queries will be used, in aggregate, to describe the breadth and depth of data in the NIH DRN and also to help inform the NIH DRN Coordinating Center regarding issues related to data availability, capture, and fitness for use. The NIH DRN Coordinating Center will create and annually update a data characterization report to help characterize NIH DRN data to assist requestors in identifying appropriate data partners for proposed studies.

4.1.2 Questionnaires and Targeted Questions

Some requests may take the form of a question or brief questionnaire (e.g., questions regarding the expected electronic capture of the provision of specific products or services, such as infused therapies or group therapy) to help identify possible data sources or expected gaps in data capture. These requests will not generate person-level information but could involve information that Data Partners consider confidential or proprietary. If the response is deemed confidential/proprietary, Data Partners should indicate so in the response, and appropriate protections should be made.

4.1.3 Analyses with Aggregate Results

Some analyses yield only aggregate data (e.g., counts, regression results), i.e., no person level data, and no cell sizes less than 6. Aggregate data can be subcategorized by demographics (e.g., sex, age group), period (e.g., month, year), and health service characteristics (e.g., comorbidity score). These analyses may be performed for project planning, proposal development, or because the results themselves are meaningful without requiring any additional analysis. Examples of requests include:

- Counts of members meeting certain criteria (e.g., women over 40 years old who have filled an antidepressant prescription)
- Counts of exposures, outcomes, or exposure/outcome pairs
- Counts of members with various health plan characteristics (e.g., enrollment months or medical benefits, pharmacy benefits)
- Rate of an outcome following exposure to a medical intervention

4.1.4 Analyses with Person-level Data

Some analyses result in the sharing of person-level data, and requestors and Data Partners will agree in advance about such activities performed as part of an NIH DRN approved request. As needed, an agreement/contract will address the level of participation by Data Partners as collaborators in the activity, the number and type of requests that will be included, response schedule, and reimbursement. Request execution will require existence of appropriate approval by IRBs and HIPAA privacy boards.

4.2 Data Request Fulfillment

The NIH DRN CC will facilitate identification of topic-appropriate data resources available within the NIH DRN, aid in developing requests, and support query distribution to implement approved research protocols. The NIH DRN operates under an “opt in” model. That is, no participating Data Partner will be expected to participate in research activities without the approval and involvement of an investigator at that Data Partner’s site. It is expected that Data Partners will engage with requestors and in activities that yield only aggregate data. Data Partner organizations and requestors are not mutually exclusive; Data Partners can also be requestors and can collaborate with other Data Partners/requestors on specific projects.

All data requests must be submitted to the NIH DRN CC using the NIH DRN request forms. The NIH DRN CC will conduct an administrative review to determine appropriateness of the query/request. This may involve clarification of the request with the requestor, and determination of appropriateness with relevant Data Partner Site PIs. In general, the NIH DRN CC will act primarily as a tracking and prioritization mechanism.

The NIH DRN CC will work with the requestor to select the most appropriate mechanism for responding to each request. Once a request is approved by the NIH DRN CC, the CC will initiate the request, manage the request process, and provide the results to the requestor.

Distributed querying is typically accomplished through the following six steps:

1) The requestor develops a question;
2) The requestor sends the question to the NIH DRN CC, using the NIH DRN request form, for distribution (via the NIH DRN portal) to potential collaborating Data Partners;
3) Each Data Partner determines whether or not to answer the request and whether or not an agreement with the requestor is needed;
4) Data Partners use their local datasets to obtain results;
5) Data Partners securely send results to the NIH DRN CC using the NIH DRN;

6) The NIH DRN CC reviews the results and submits the response to the requestor. Results are often aggregate results, without confidential or proprietary data. The level of data sharing is determined, in advance, as part of the collaboration agreement, and all query responses can be reviewed by the Data Partner before they are released.

4.2.1 Standard Queries

Standard queries refer to a series of specific query types/approaches that are supported by the NIH DRN CC and that can be pre-approved by the Data Partners as an acceptable form for standard querying. These query types have several important characteristics, including 1) the use of standardized query forms, 2) generation of query results in standardized formats, 3) use of an approved query code and system tested and maintained by the NIH DRN CC, and 4) use of a query approach tested and accepted by Data Partners. Data Partners will only receive query requests that use query types that they have agreed to accept.

There are currently three types of "standard queries" supported by the NIH DRN CC (listed below). Data Partners must identify which of these query types they are willing and able to receive and respond to quickly. Additional standard query types will be added as new functionality is developed and new models are adopted by Data Partners.

4.2.1.1 Summary Table Queries

The NIH DRN can query pre-compiled tables created and maintained by the FDA Mini-Sentinel program. These include nine types of prevalent conditions or treatment queries, three types of incident conditions or treatments queries, and several "Most Frequent Utilization" queries.

4.2.1.2 Modular Program Requests

The NIH DRN can use standardized, modular SAS programs that allow requesters to execute standardized analyses using individually selected parameters (e.g., exposures, outcomes, query period, age groups, etc.) that can be modified for each request. The programs themselves are not modified, only the parameter file used in the request changes. Seven modular programs have been developed and published by the FDA Mini-Sentinel program; these can be executed against the Mini-Sentinel Common Data Model. Additional modular programs may be created by the NIH DRN to perform additional analyses and/or to operate on different data models.

4.2.1.3 EHR Support for Public Health (ESP) Queries

The NIH DRN can use a simple query interface, the ESPnet Query Builder, to query data formatted according to the ESPnet data model. These queries can execute against the ESP data model.

4.2.2 Customized Queries

Customized programs may be necessary to address questions not answerable with standard queries, either because the question requires different kinds of analysis, or to work with a data model that the standard queries do not support. These programs may be developed by the NIH DRN CC, in collaboration with the requestor, or by the requestor, if the requestor has expertise in programming in the NIH DRN data environment. As with all requests, the NIH DRN CC must approve the allocation of resources to develop the query to fulfill a data request. Once approved, the query will be developed and tested by the NIH DRN CC. The NIH DRN portal will be used to distribute these requests, and to return results. Data Partners will be notified about the development of customized queries and will participate on an opt-in and case-by-case basis.

4.2.3 Meta Data Requests and Surveys

Some requests do not require access to data, but rather focus on health plan characteristics or other non-data information. The NIH DRN CC will work with the requestor to assess the viability and preferred approach for collecting this type of information. Examples include formulary status changes, insurance product offerings, and population characteristics.

4.3 Security Policies

The NIH DRN will comply with all applicable federal, state and professional standards, including those promulgated by the HIPAA and NIH.

The following list contains major security policies of the system:

- Users are required to select strong passwords with the following rules at minimum: at least eight characters, at least one number, at least one nonnumeric character, at least one capital letter, at least one lower case letter. Passwords cannot contain the username or any part of the Network participant's full name.
- Passwords will be changed at least every six months
- Passwords cannot be re-used
- Computers will be automatically locked down or logged off at maximum after thirty minutes of inactivity
- The NIH DRN System Administrators will verify NIH DRN participants' identities and email addresses before creating new user accounts
- Users must use corporate email addresses for NIH DRN communication
- All NIH DRN activity will be logged (a record will be kept of access, user identification [ID] changes, query initiation, results upload, etc.)
- NIH DRN CC will regularly review audit logs
- Antivirus software will run regularly on all NIH DRN system servers
- Data Partner analysts will be able to create audit logs of all of their NIH DRN activity
- System communication will use Hypertext Transfer Protocol Secure/Secure Socket Layer (HTTPs/SSL/TLS) standards
- The NIH DRN will be hosted in a secure, FISMA compliant data center

5. Communications

Query requests should be initiated as described above (section 4.1). Questions or comments to the NIH DRN Governance Advisory Committee can be addressed to Beth Syat (beth_syat@harvardpilgrim.org).

The NIH DRN Governance Advisory Committee will meet regularly and disseminate information through regular updates to the HCS Collaboratory Coordinating Center and on Steering Committee calls.

5.1 Dissemination of Results

The EHR Core will regularly publish descriptive statistics and quality assurance measures on the public portal, with the permission of each Data Partner PI and the HCS Collaboratory Coordinating Center.

6. Confidentiality

Protecting confidentiality of data is a critical component of the NIH DRN. This governance document pertains to the work of the EHR Core only. Each query/request will have the appropriate and necessary human subjects approvals and follow confidentiality procedures. The NIH DRN Coordinating Center will request only the minimum necessary information and use aggregate data first to answer queries. Under no circumstances will identifiers, as defined by HIPAA, be included in a data set. Identifiers include name, Social Security Number, medical record number, and address (excluding ZIP Code). Numbering schemes that cannot be associated with any of the above identifiers may be introduced for linking data from individuals.

Attachment A

The purpose of this document is to serve as a placeholder for describing specific information about query capacity of the NIH DRN's Coordinating Center and data partners. This is a living document that will be updated to reflect currently-agreed upon terms.

As of [date]:

- **The NIH DRN Coordinating Center** will process up to X modular program requests and up to X summary table requests or other menu-driven, rapid-response requests per month, as directed by NIH.

- **Data Partners** will respond to:
    - Summary table requests or other menu-driven rapid-response requests: process up to X per month and respond within X working days or provide a reason for not responding
    - Modular program requests: process up to X per month and provide one of the following responses within X working days of receipt of a program:
        - Accept query and run it within allotted timeframe;
        - Ask for additional information; or
        - Reject query and provide a reason