Distributed Research Network: Policies and Procedures

Table of Contents

1. Introduction

1.1. Purpose of This Document

This document provides an overview of the National Institutes of Health (NIH) Pragmatic Trials Collaboratory's Distributed Research Network (DRN) governance policies, querying capabilities, and responsibilities. These policies apply to NIH Collaboratory 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 within the NIH Collaboratory. NIH Collaboratory DRN activities are supported by the 1U24AT009676-01 cooperative agreement; and manuscripts, abstracts, and presentations derived from DRN activities are subject to the NIH Collaboratory Publications and Presentations Policy.**

1.2. Background

The NIH Pragmatic Trials Collaboratory DRN leadership team at the Harvard Pilgrim Health Care Institute (HPHCI) is responsible for organizing and leading the DRN, with input from site principal investigators (PIs) and oversight by the NIH Collaboratory Coordinating Center PIs. The NIH Collaboratory DRN is described in detail at rethinkingclinicaltrials.org/nih-collaboratory-drn.

In brief, the NIH Pragmatic Trials Collaboratory DRN is a collaborative enterprise that facilitates research partnerships with organizations that participate in the FDA's Sentinel Program, using separate governance, oversight, and funding. Where appropriate, the DRN uses data sets and querying methods developed by the Sentinel Program. The network is administered by the NIH Collaboratory DRN Coordinating Center, which also provides query support and other services to facilitate use of the network.

The NIH Pragmatic Trials Collaboratory DRN can facilitate research partnerships by connecting investigators funded by the NIH and other not-for-profit sponsors to DRN data partners (organizations that maintain electronic health data) to identify opportunities for collaboration. The network is best used to facilitate multisite research.

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

2. Roles and Responsibilities

2.1. DRN Coordinating Center

The NIH Pragmatic Trials Collaboratory DRN Coordinating Center resides at the Harvard Pilgrim Health Care Institute (HPHCI) and is responsible for the day-to-day operations of the DRN. These responsibilities include:
- Query Support
  - Develop and oversee the query request, approval, and response process.

- o Support NIH staff and other requesters regarding use of the DRN, querying capabilities, data resources, and interpretation of results. Support also includes data fitness for use and epidemiologic expertise.
  - o Ensure that all data requests are tested and appropriate for use within the DRN and with the specific data partners involved (see Query Capacity and Fulfillment).
- Query Capacity and Fulfillment
  - o Create the capability and expertise to use publicly available tools, including, but not limited to, those developed by the FDA Sentinel Initiative (www.sentinelinitiative.org), and the existing querying features of the open source version of PopMedNet (www.popmednet.org). Tools available include the Sentinel modular programs and other custom programs.
  - o Work with data partners to establish their engagement and agreements/contracts and assist with query response.
  - o The DRN Coordinating Center will establish a response and reimbursement agreement with data partners regarding the process for approving and responding to queries. Agreements will address:
  - o Procedures that ensure data partners' ability to review and approve queries and results
  - o Expectations regarding timeliness of responses Compensation
  - o Review and editorial rights of reports and publications, including but not limited to data provided, shown, and interpretations
  - o Authorship and/or acknowledgement by investigators on publications and reports
- Network Maintenance and Software
  - o Incorporate software updates and bug fixes
  - o Secure hosting and maintenance
- Additional Activities
  - o Collaborate with the NIH Collaboratory's Data Sharing Committee to help implement uniform data sharing principles and use of analytic data sets.
  - o Request information from data partners related to data availability and fitness for use.
  - o Write distributed programs to be run at data partner sites, for both quality assurance and approved queries.
  - o Approve, prioritize, test, and execute approved requests.
  - o Develop and oversee the query request and response process.
  - o Ensure that documentation of data partners' site characteristics, data availability, and quality are available internally and publicly.
  - o Ensure all data queries/requests have a protocol/research proposal and appropriate institutional review board (IRB) approval and are submitted by an approved requestor.
  - o 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.

2.2. Data Partners

Data partners are member organizations that have electronic health data resources in Sentinel Common Data Model format and the technical capacity to respond to a query via the NIH Pragmatic Trials Collaboratory DRN. Some data partners have the ability to respond to rapid-response queries. The DRN Coordinating Center will only send requests to data partners that have self-identified as having the ability and willingness to respond to such requests.

Each 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 data set. 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.

2.2.1. 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 set(s)
- Meet all of the compliance requirements applicable to their local sites
- Assign 1 or more local analysts to receive and respond to queries/requests
- Respond to all queries/requests per agreed upon guidelines
- Maintain a list of its current staff who are authorized to participate in the NIH Pragmatic Trials Collaboratory DRN, their contact information, and their roles and responsibilities within the DRN
- Obtain and maintain local approval related to participation in NIH Collaboratory 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

2.2.2. Data Partner PIs

The NIH Pragmatic Trials Collaboratory 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 DRN Coordinating Center queries/requests.

- Communicate clearly and in a timely manner to the DRN Coordinating Center if they choose to "opt out" of an activity.

2.2.3. Data Partner Project Managers

Data partner project managers are expected to:
- Liaise between the DRN Coordinating Center and their sites
- Communicate with their local site PIs to ensure applicable local site compliance requirements are met

2.2.4. Data Partner Analysts

Data partner analysts will:
- Respond to queries/requests from the DRN Coordinating Center within the specified timeframe, as directed by the site PI or local approval workflow
- Provide documentation to the DRN Coordinating Center on site-specific data issues
- Follow established procedures for data queries/requests
- Attend calls and meetings, as needed

2.3. Requestors

Authorized requestors include NIH Pragmatic Trials Collaboratory leadership, 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 DRN Coordinating Center. If approved by the DRN Coordinating Center, all queries/requests will be submitted by the DRN Coordinating Center staff to the data partners through the DRN's secure query tool.

Requestors are responsible for:
- 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 DRN Coordinating Center and the NIH
- Collaboratory's EHR Core Working Group
- Obtaining necessary approvals by IRBs and Health Insurance Portability and Accountability Act (HIPAA) privacy boards
- Working with the DRN Coordinating Center and the data partners to execute necessary agreements/contracts

Requestors are expected to:
- Use results provided only for the stated and approved purpose, which may include a requirement to keep results confidential
- Abide by any other limitations on use, issued by the NIH Pragmatic Trials Collaboratory's EHR Core Working Group, the DRN Coordinating Center, and/or the data partner providing and reporting data

3. Policies and Procedures

3.1. Data Queries and Requests

The procedures outlined below are designed to balance efficiency and data partner autonomy. The DRN Coordinating Center can help determine which kind of query/request is most appropriate for each situation.

For each query/request, there are 4 steps: initiation, approval, execution, and reporting. Regardless of the approach, the following information will be required from the requestor(s): details of the request, purpose of the request, intended use of the results, expected level of protected health information in the response, and the date by which the information is needed.

3.1.1. Data Completeness and Data Characterization Activities

The DRN Coordinating Center will issue queries/requests to ensure proper system functionality and to help characterize data available within the 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 DRN and to help inform the DRN Coordinating Center regarding issues related to data availability, capture, and fitness for use. The DRN Coordinating Center will create and annually update a data characterization report to help characterize DRN data to assist requestors in identifying appropriate data partners for proposed studies.

3.1.2. 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.

3.1.3. Analyses and Aggregate Results

Some analyses yield only aggregate data—such as counts or regression results—that is, no person-level data and no cell sizes less than 6. Aggregate data can be subcategorized by demographic characteristics (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 additional analysis. Examples of requests include:
- Counts of members meeting certain criteria (e.g., women older than 40 years 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

3.1.4. Analyses With Person-Level Data

Some analyses may result in the sharing of person-level data, and requestors and data partners will agree in advance about such activities performed as part of a 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.

3.2. Data Request Fulfillment

The DRN Coordinating Center will facilitate identification of topic-appropriate data resources available within the DRN, aid in developing requests, and support query distribution to implement approved research protocols. The 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 DRN Coordinating Center using the DRN request forms. The DRN Coordinating Center will conduct an administrative review to determine the 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 DRN Coordinating Center will act primarily as a tracking and prioritization mechanism.

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

Distributed querying is typically accomplished through the following 6 steps:
1. The requestor develops a question;
2. The requestor sends the question to the DRN Coordinating Center, using the DRN request form, for distribution (via the 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 data sets to obtain results;
5. Data partners securely send results to the DRN Coordinating Center using the DRN;
6. the DRN Coordinating Center 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.

3.2.1. Standard Queries

Standard queries refer to queries that use Sentinel's modular programs. 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 DRN Coordinating Center, and (4) use of a query approach tested and accepted by data partners. Data partners will only receive query requests that they have agreed to accept.

Modular programs allow requesters to execute standardized analyses using individually selected parameters (e.g., exposures, outcomes, query period, age groups) that can be modified for each request. The programs themselves are not modified; only the parameter file used in the request changes.

3.2.2. 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 DRN Coordinating Center, in collaboration with the requestor, or by the requestor, if the requestor has expertise in programming in the DRN data environment. Once approved, the query will be developed and tested by the DRN Coordinating Center. The 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.

3.2.3. Metadata Requests and Surveys

Some requests do not require access to data, but rather focus on health plan characteristics or other non-data information. The DRN Coordinating Center 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.

3.3. Security Policies

The NIH Pragmatic Trials Collaboratory DRN will comply with all applicable federal, state, and professional standards, including those promulgated by the HIPAA and the 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 8 characters, at least 1 number, at least 1 nonnumeric character, at least 1 capital letter, at least 1 lower case letter. Passwords cannot contain the username or any part of the DRN participant's full name.
- Passwords will be changed at least every 6 months.
- Passwords cannot be reused.
- Computers will be automatically locked down or logged off at maximum after 30 minutes of inactivity.
- The DRN system administrators will verify DRN participants' identities and email addresses before creating new user accounts.
- Users must use corporate email addresses for DRN communication.

- All DRN activity will be logged (a record will be kept of access, user identification changes, query initiation, results upload, etc.).
- The DRN Coordinating Center will regularly review audit logs.
- Antivirus software will run regularly on all DRN system servers.
- Data partner analysts will be able to create audit logs of all of their DRN activity.
- System communication will use hypertext transfer protocol secure/secure socket layer/transport layer security (HTTPS/SSL/TLS) standards
- The DRN will be hosted in a secure, FISMA-compliant data center.

4. Communications

Query requests should be initiated as described above (section 3.1). Questions or comments to the DRN Governance Advisory Committee can be addressed to Audrea Wolfe (audrey_wolfe@harvardpilgrim.org).

5. Confidentiality

Protecting confidentiality of data is a critical component of the NIH Pragmatic Trials Collaboratory DRN. Each query/request will have the appropriate and necessary human subjects approvals and follow confidentiality procedures. The 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.