

Use and Sustainability of the Collaboratory Distributed Research Network

Lesley Curtis, PhD & Richard Platt, MD, MSc
for the DRN Team

What is the Collaboratory DRN trying to accomplish?

Goal: Facilitate multisite research collaborations between investigators and data partners by creating secure networking capabilities and analysis tools.

- Advantages of a distributed research network (DRN)
 - ✓ Ability to work with analysis-ready datasets covering many millions
 - ✓ Standardized data using a common data model
 - ✓ Data partners keep and analyze their own data
 - ✓ Provide results, not data, to the requestor
 - ✓ All activities audited and secure

NIH Collaboratory Distributed Research Network Partners

■ NIH Collaboratory Distributed Research Network

Millions of people. Strong collaborations. Privacy first.

Data Partners



The **Meyers** Primary Care Institute

All participate in FDA's Sentinel System



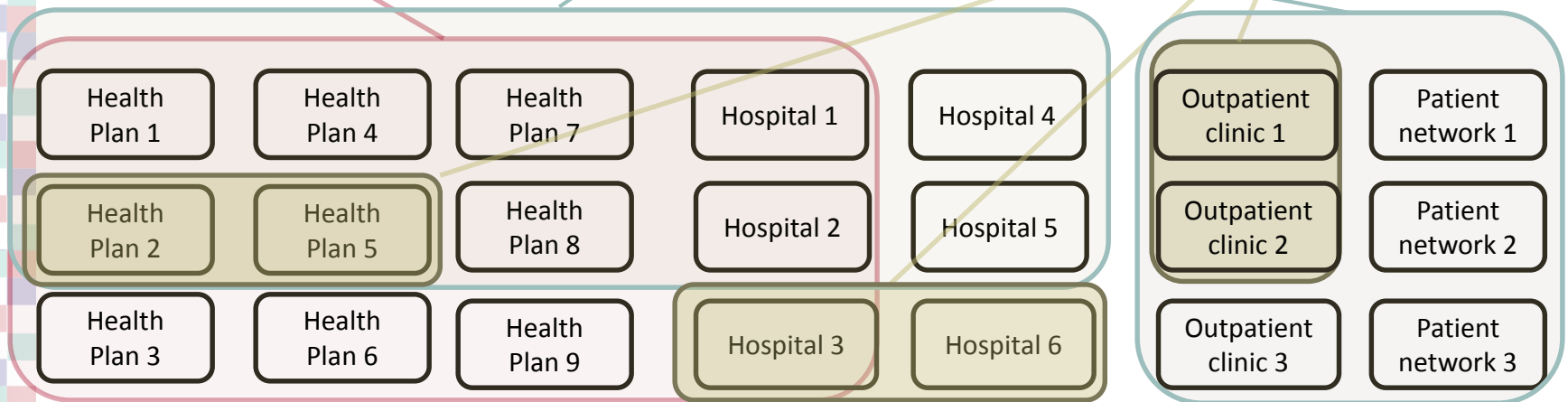
Available data

- Rapid-response distributed querying available across data partners with over 90 million lives
- >300 million person-years of observation time
- Detailed information for billions of medical encounters and outpatient pharmacy dispensings
- Analysis-ready datasets (i.e., quality checked and formatted) representing >90% of the FDA Sentinel program

How can the DRN be used?

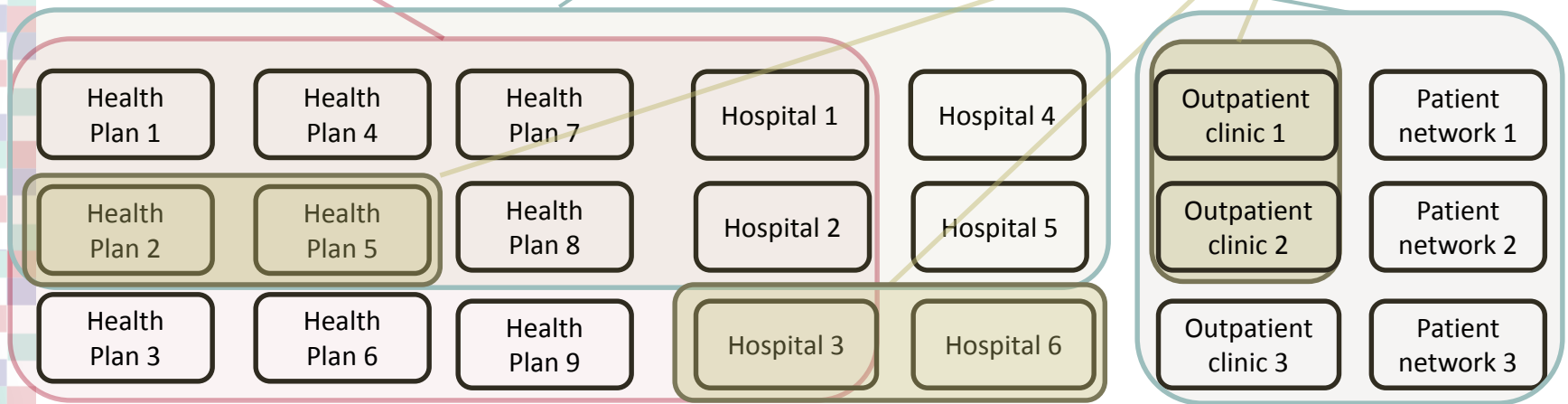
- Research planning
 - Assess disease burden/outcomes
 - Prioritize research domains
 - Pragmatic clinical trial design
- Answer a question!
- Pragmatic clinical trial follow-up
- Platform for data sharing and reuse

Critical partners in the national infrastructure



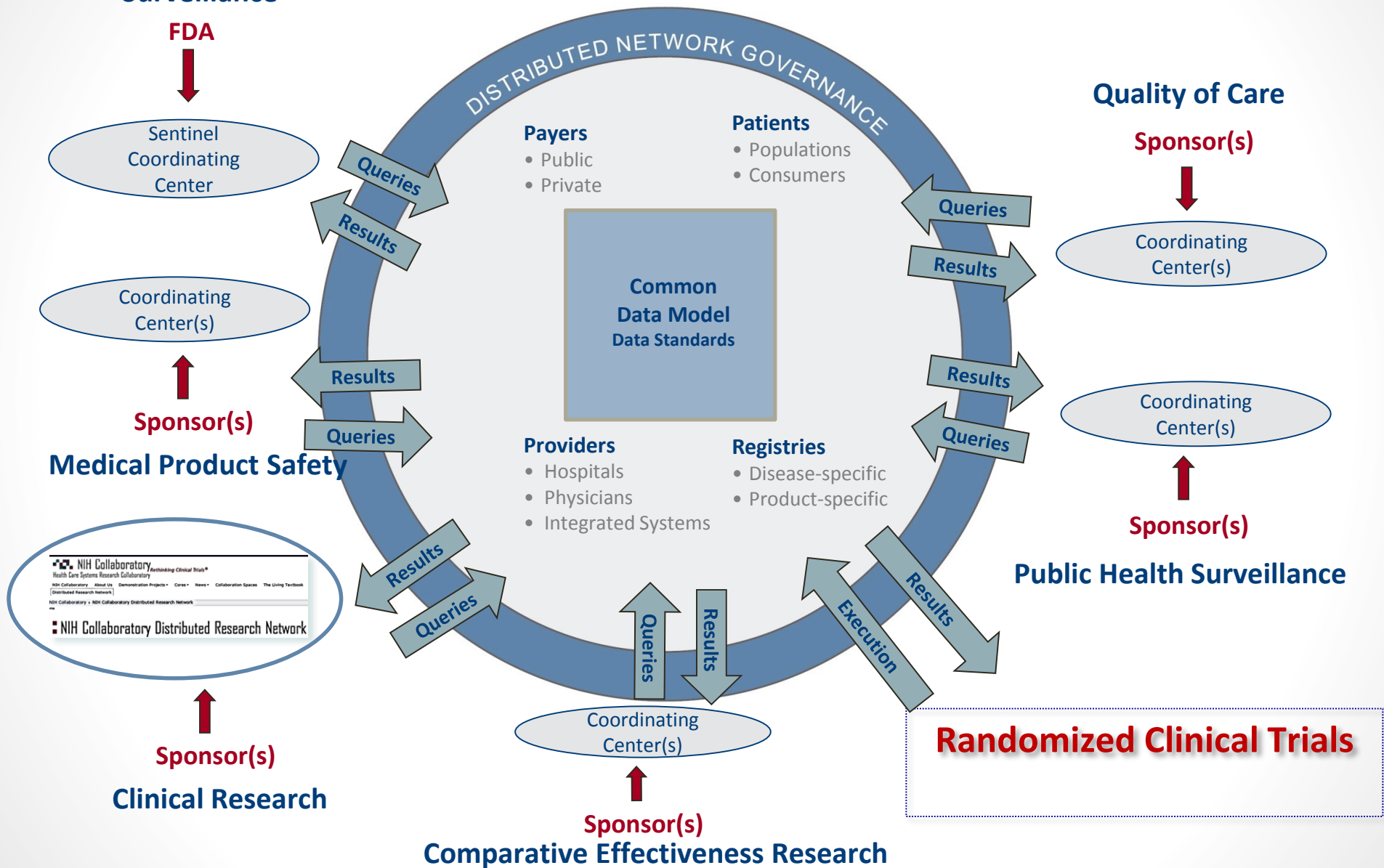
- Each organization can participate in multiple networks
- Each network controls its governance and coordination
- Networks share infrastructure, data curation, analytics, lessons, security, software development

Critical partners in the national infrastructure



- Each organization can participate in multiple networks
- Each network controls its governance and coordination
- Networks share infrastructure, data curation, analytics, lessons, security, software development

Medical Product Safety National Evidence Generation Concept





The Collaboratory DRN and PCORnet are complementary

- Collaboratory DRN is based on administrative claims and outpatient pharmacy dispensing data
 - Complete data for most reimbursed care → if no evidence of an event, it very likely didn't occur
 - Limited access to medical record information
- PCORnet is based on EHR data
 - Detailed information care provided by clinical organization, including vital signs, lab test results
 - Limited information about care provided by other organizations or drug dispensing



Moving toward a sustainable DRN

- Build the Collaboratory DRN user community
 - NIH Institutes and Centers
 - NIH Investigators
- Develop of data sharing resources



Moving toward a sustainable DRN

- **Build the Collaboratory DRN user community**
 - **NIH Institutes and Centers**
 - **NIH Investigators**
- **Develop data sharing resources**



Build the user community within NIH institutes and centers

- Special NIH supplement in 2014 for pilot test using publicly available Sentinel querying tools
- Three pilot test queries developed with NCCAM, NHLBI, NCI
 - Assess recruitment feasibility of replicating the Trial to Assess Chelation Therapy (TACT)
 - Characterize statin users >75 years of age
 - Assess rates of abnormal cancer screening test results and rates of follow up testing
- Queries used as test cases for developing processes and refining strategies to develop queries



Learnings from the pilot

- Data Partners will respond
 - Participation from 3 data partners, representing $\sim 1/3$ of the total data
 - Participation largely a function of available funds
- Translating a question to an answerable query often requires an iterative process
 - Some questions may be readily addressed with a simple query
- User guides and training materials are essential

<https://www.nihcollaboratory.org/Pages/distributed-research-network.aspx>

The screenshot shows a Firefox browser window displaying the NIH Collaboratory Distributed Research Network website. The address bar shows the URL <https://www.nihcollaboratory.org/Pages/distributed-research-network.aspx>. The website header features the NIH Collaboratory logo with the tagline "Rethinking Clinical Trials®" and a search bar. A navigation menu includes links to "NIH Collaboratory", "About Us", "Demonstration Projects", "Cores", "News", "Collaboration Spaces", "Knowledge Repository", and "The Living Textbook". Below the navigation menu, a breadcrumb trail reads "NIH Collaboratory > NIH Collaboratory Distributed Research Network". The main heading is "NIH Collaboratory Distributed Research Network". The subheading reads "Millions of people. Strong collaborations. Privacy first." The main text describes the network's purpose: "The NIH Collaboratory Distributed Research Network enables investigators to collaborate with each other in the use of electronic health data, while also safeguarding protected health information and proprietary data. It supports both single- and multisite research programs. The Network's querying capabilities reduce the need to share confidential or proprietary data by enabling authorized researchers to send queries to collaborators holding data (i.e., data partners). In some cases, queries can take the form of computer programs that a data partner can execute on a preexisting dataset. The data partner can return the query result, typically aggregated (count) data, rather than the data itself. This form of remote querying reduces legal, regulatory, privacy, proprietary, and technical barriers associated with data sharing for research. The network seeks to build strong and trusted collaborations to support the research that will lead to improved health for millions of people around the world." A section titled "What does the NIH Collaboratory Distributed Research Network do?" lists five bullet points: "Provides infrastructure and mechanisms to facilitate multicenter studies using electronic clinical, administrative, and research data", "Allows searchable discovery of available data resources, health systems, researchers, and re-usable analytic tools", "Enables authorized investigators to identify clinical, administrative, and research datasets of interest", "Facilitates multisite distributed querying of data resources, while allowing the data to remain in the control of the data owners", and "Serves as a repository of tools to leverage EHRs to support clinical research across multiple health systems". A sidebar on the right contains three links: "DRN Governance Document, v1.0", "Distributed Research Network User's Guide, v5.0 (PDF)", and "DRN Request Form (.docx)". A box on the left side of the main content area contains the text "To learn more about the NIH Collaboratory Distributed Research Network" and the email address "support@popmednet.org". The browser's taskbar at the bottom shows the Start button and several open applications, including Outlook 2013, NIH Collaboratory Dis..., and Collaboratory DRN_June... The system clock in the bottom right corner shows 11:25 AM.

Firefox NIH Collaboratory Distributed Research Net...
https://www.nihcollaboratory.org/Pages/distributed-research-network.aspx
Most Visited Getting Started From Internet Explorer Harvard Pilgrim Health... Resource Center | PC...
Search this site...
NIH Collaboratory Rethinking Clinical Trials®
Health Care Systems Research Collaboratory
NIH Collaboratory About Us Demonstration Projects Cores News Collaboration Spaces Knowledge Repository The Living Textbook
Grand Rounds
NIH Collaboratory > NIH Collaboratory Distributed Research Network
NIH Collaboratory Distributed Research Network
Millions of people. Strong collaborations. Privacy first.
The NIH Collaboratory Distributed Research Network enables investigators to collaborate with each other in the use of electronic health data, while also safeguarding protected health information and proprietary data. It supports both single- and multisite research programs.
The Network's querying capabilities reduce the need to share confidential or proprietary data by enabling authorized researchers to send queries to collaborators holding data (i.e., data partners). In some cases, queries can take the form of computer programs that a data partner can execute on a preexisting dataset. The data partner can return the query result, typically aggregated (count) data, rather than the data itself. This form of remote querying reduces legal, regulatory, privacy, proprietary, and technical barriers associated with data sharing for research.
The network seeks to build strong and trusted collaborations to support the research that will lead to improved health for millions of people around the world.
What does the NIH Collaboratory Distributed Research Network do?
• Provides infrastructure and mechanisms to facilitate multicenter studies using electronic clinical, administrative, and research data
• Allows searchable discovery of available data resources, health systems, researchers, and re-usable analytic tools
• Enables authorized investigators to identify clinical, administrative, and research datasets of interest
• Facilitates multisite distributed querying of data resources, while allowing the data to remain in the control of the data owners
• Serves as a repository of tools to leverage EHRs to support clinical research across multiple health systems
To learn more about the NIH Collaboratory Distributed Research Network
support@popmednet.org
DRN Governance Document, v1.0
Distributed Research Network User's Guide, v5.0 (PDF)
DRN Request Form (.docx)
Start Outlook 2013 NIH Collaboratory Dis... Collaboratory DRN_June... 11:25 AM



Build the user community of NIH investigators

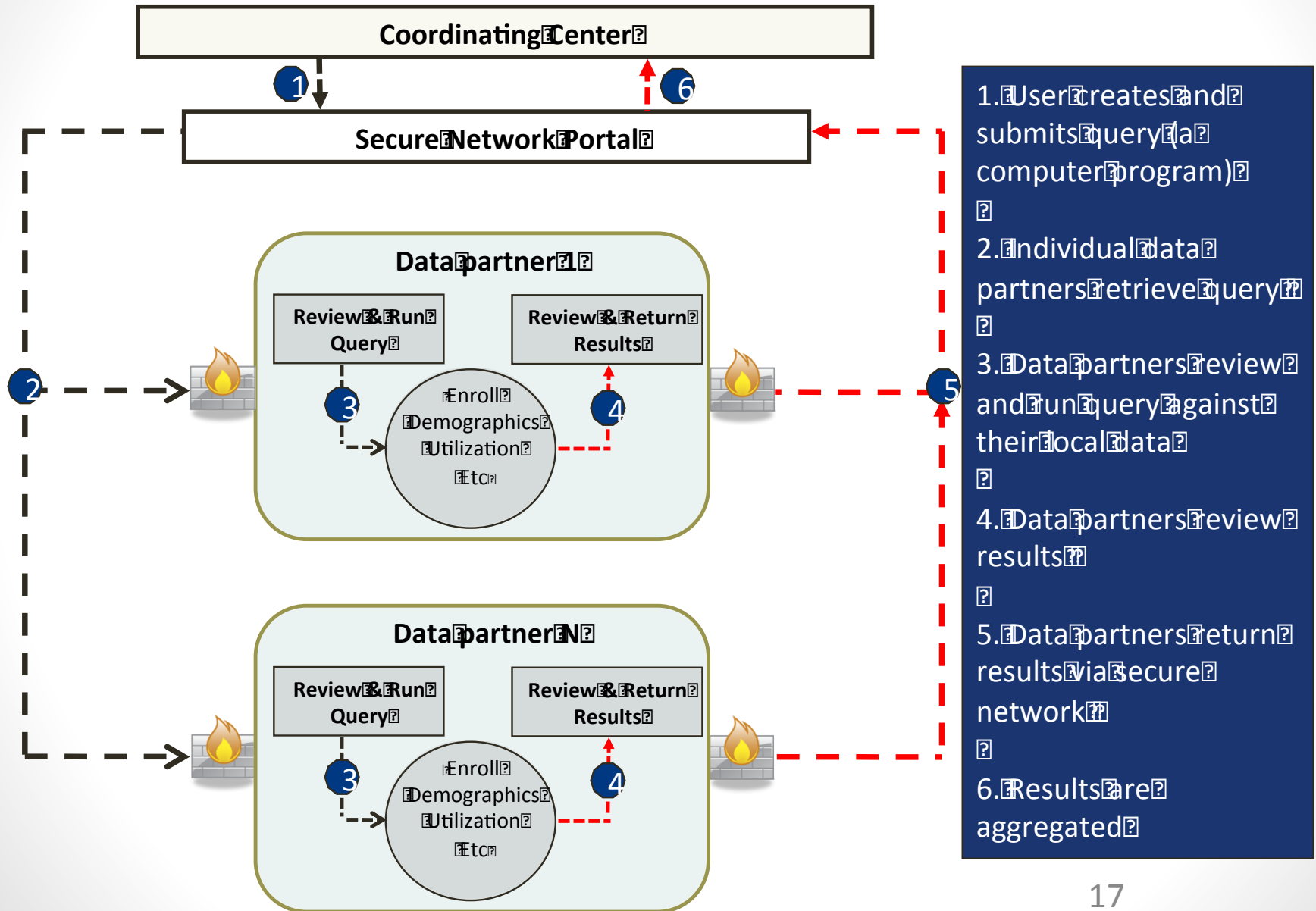
- Allocate existing resources toward pre-research queries posed by researchers
- Sponsor a mini-competition with a transparent selection process
 - Announce on Grand Rounds and the Collaboratory website
 - Establish selection criteria
 - Appropriateness of question for the DRN
 - Investigator track record
 - Potential fundability of research concept
- Use publicly available analysis tools
- Refine user guide and training materials
- Develop pricing model



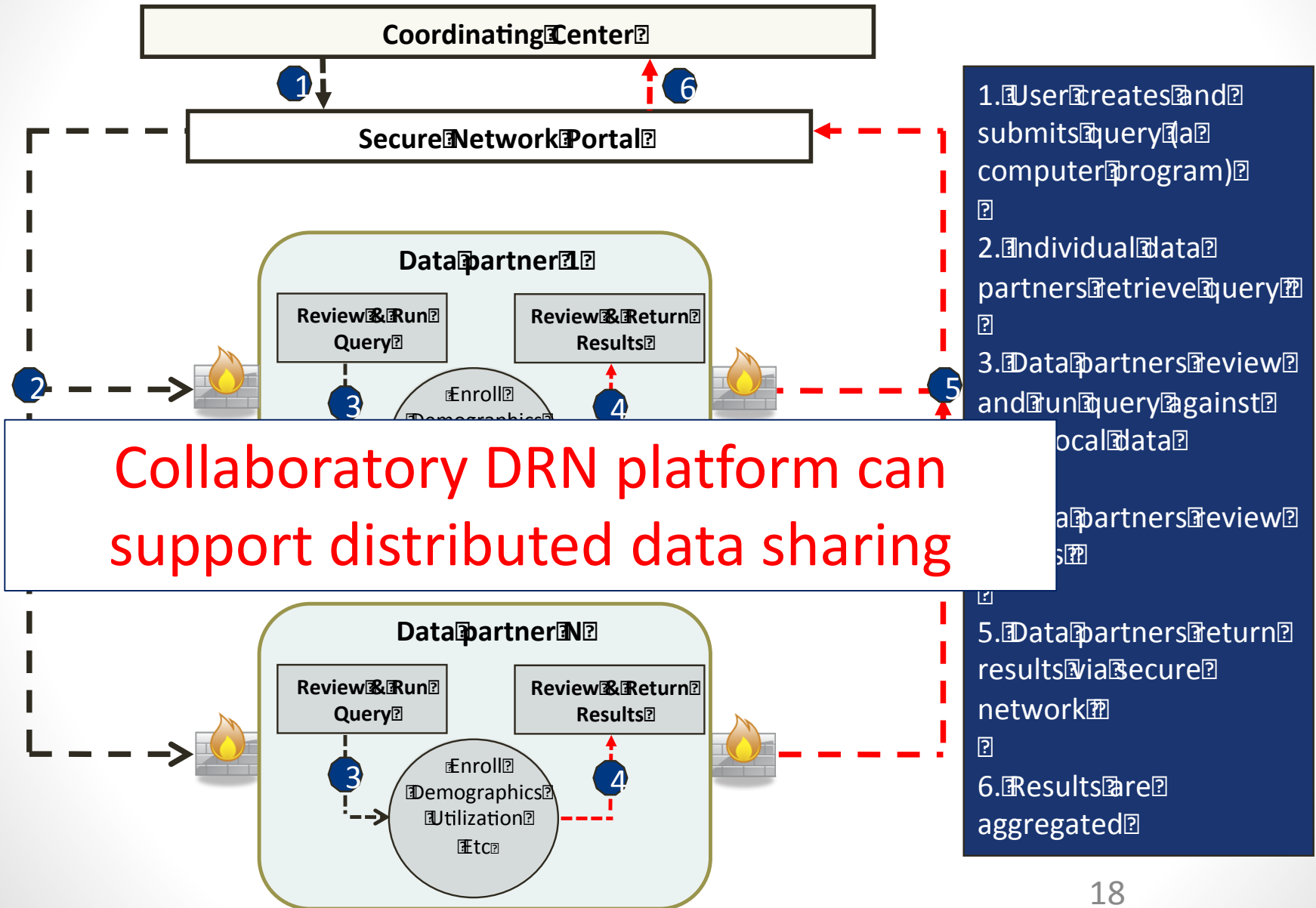
Moving toward a sustainable DRN

- Build the Collaboratory DRN user community
 - NIH Institutes and Centers
 - NIH Investigators
- **Develop data sharing resources**

What is a distributed research network?



What is a distributed research network?





Collaboratory DRN platform can support distributed data sharing

- Enables re-use of research dataset with appropriate controls for patient privacy, access, governance, and proprietary concerns
- Distributed analyses limited to the software/hardware capabilities of the enclave
- Governance over usage must be established and implemented for each resource

Key elements of a distributed data sharing platform

- Discovery of available data resources and organizations
- Information about data use requirements
- Query interface and distribution
- Access controls and permissions
- Secure and auditable infrastructure
- Data storage and governance (if investigators do not want to maintain local control of data source)

Develop resources to support data sharing

- Description of access and use restrictions
- Data dictionaries, documentation, analytic code
- Listing of publications based on data
- Tools available for use with the dataset
- Availability of test data
- Contact information



Additional Information

- For additional information, please go to:
<https://www.nihcollaboratory.org/Pages/distributed-research-network.aspx#HowSubmit>