

Uses of the NIH Collaboratory Distributed Research Network

Jeffrey Brown, PhD for the DRN Team
Harvard Pilgrim Health Care Institute and
Harvard Medical School
March 11, 2016



The Goal

The NIH Collaboratory DRN facilitates research partnerships with organizations (Data Partners) that possess **electronic health data that have been quality checked and formatted** to support multi-site biomedical research

Collaboratory DRN Objective

Goal: Facilitate multisite research collaborations between investigators and data stewards 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 stewards keep and analyze their own data
 - ✓ Provide results, not data, to the requestor
 - ✓ All activities audited and secure



Uses of the Distributed Network

- Research planning
 - Assess background rates and population impact of conditions / treatments
 - Prioritize research domains
 - Identify sites for participation in interventional or observational studies
- Answer research questions!

Requestors do not have to be experts in use of healthcare data

- Coordinating Center helps requestors understand and use the network
 - Assess fit between requests and DRN capabilities
 - Suggest ways to maximize usefulness of DRN data and resources

NIH Collaboratory Distributed Research Network Partners

■ NIH Collaboratory Distributed Research Network

Millions of people. Strong collaborations. Privacy first.

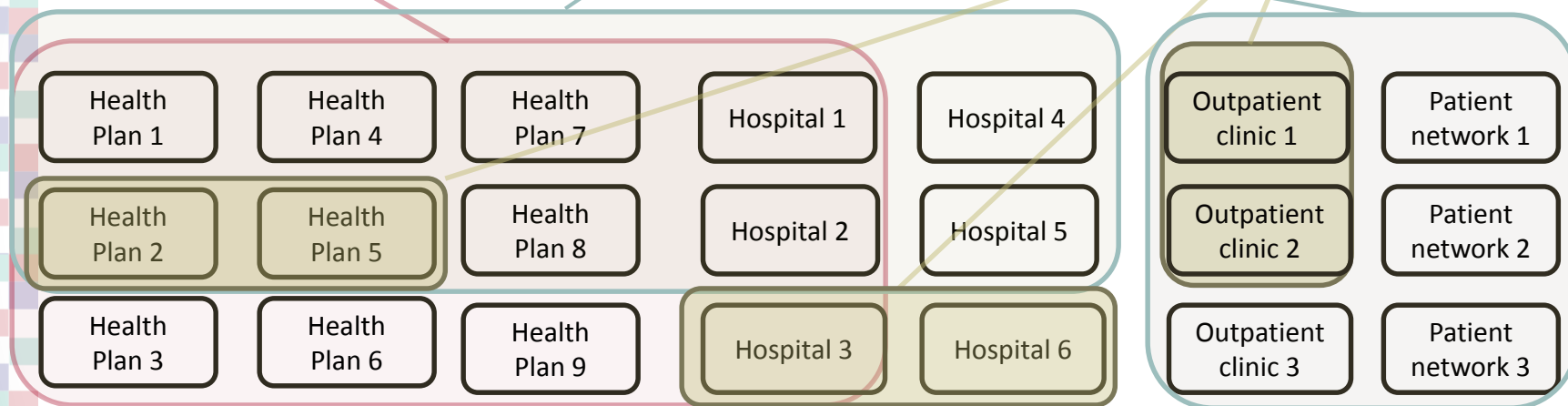
Data Partners



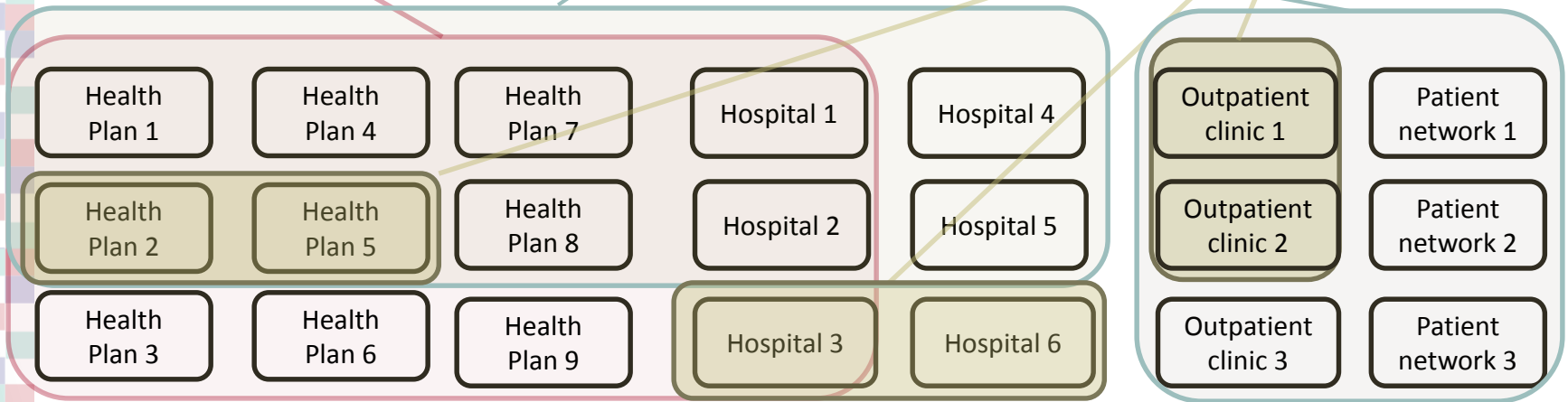
The **Meyers** Primary Care Institute

All participate in FDA's Sentinel System

Critical Partners in a National Infrastructure

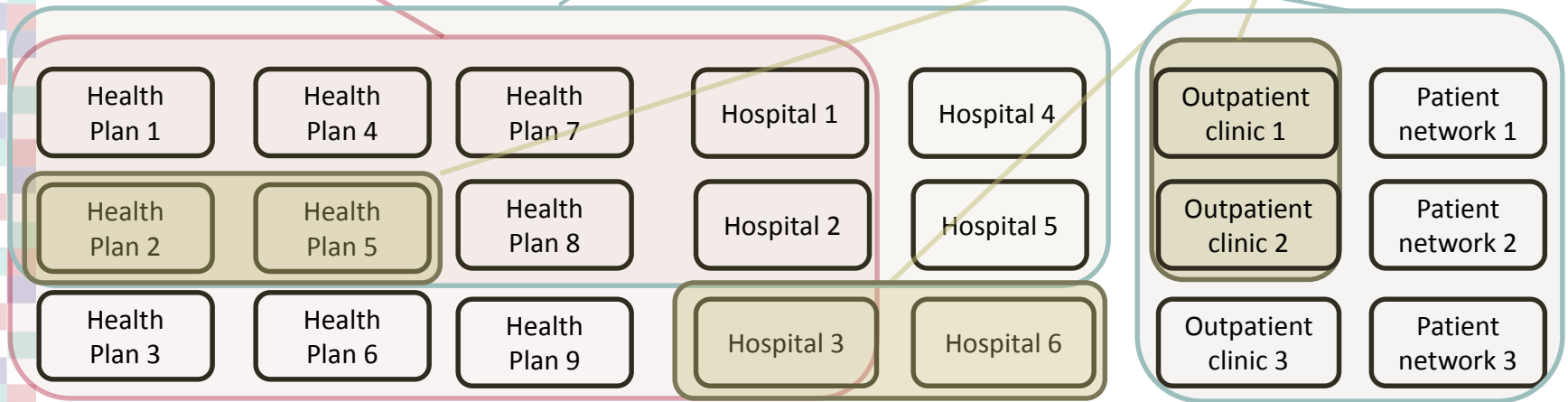


Critical Partners in a National Infrastructure



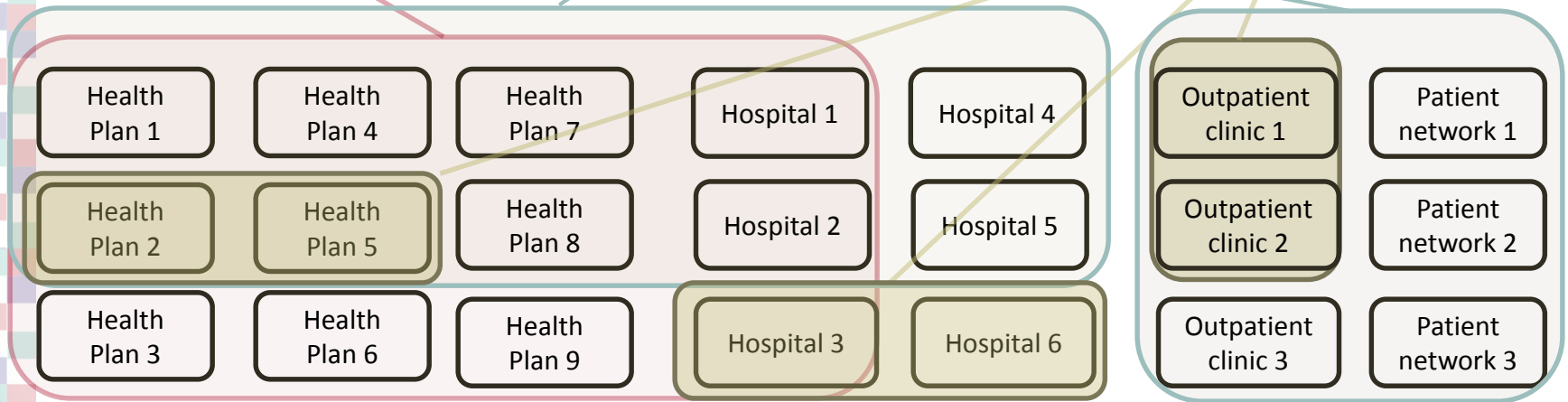
- Each organization can participate in multiple networks

Critical Partners in a National Infrastructure



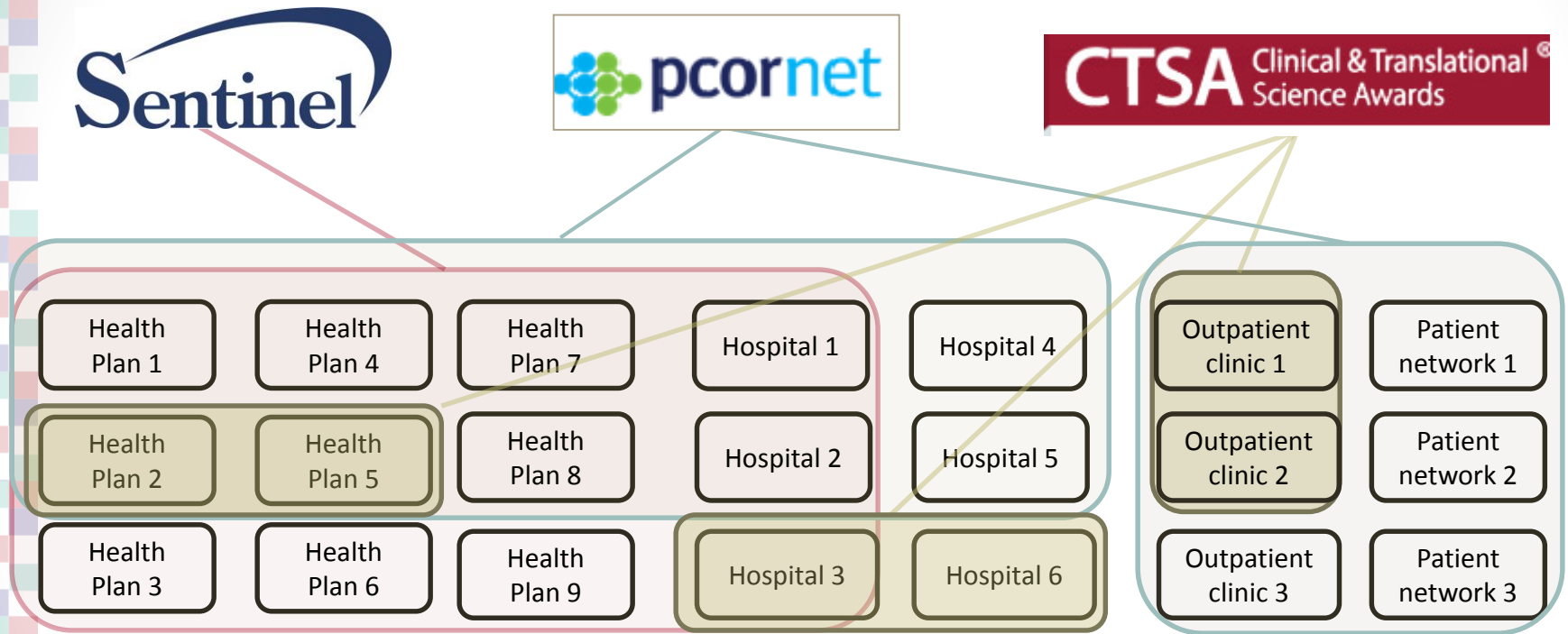
- Each organization can participate in multiple networks
- Each network controls its governance and coordination

Critical Partners in a National Infrastructure



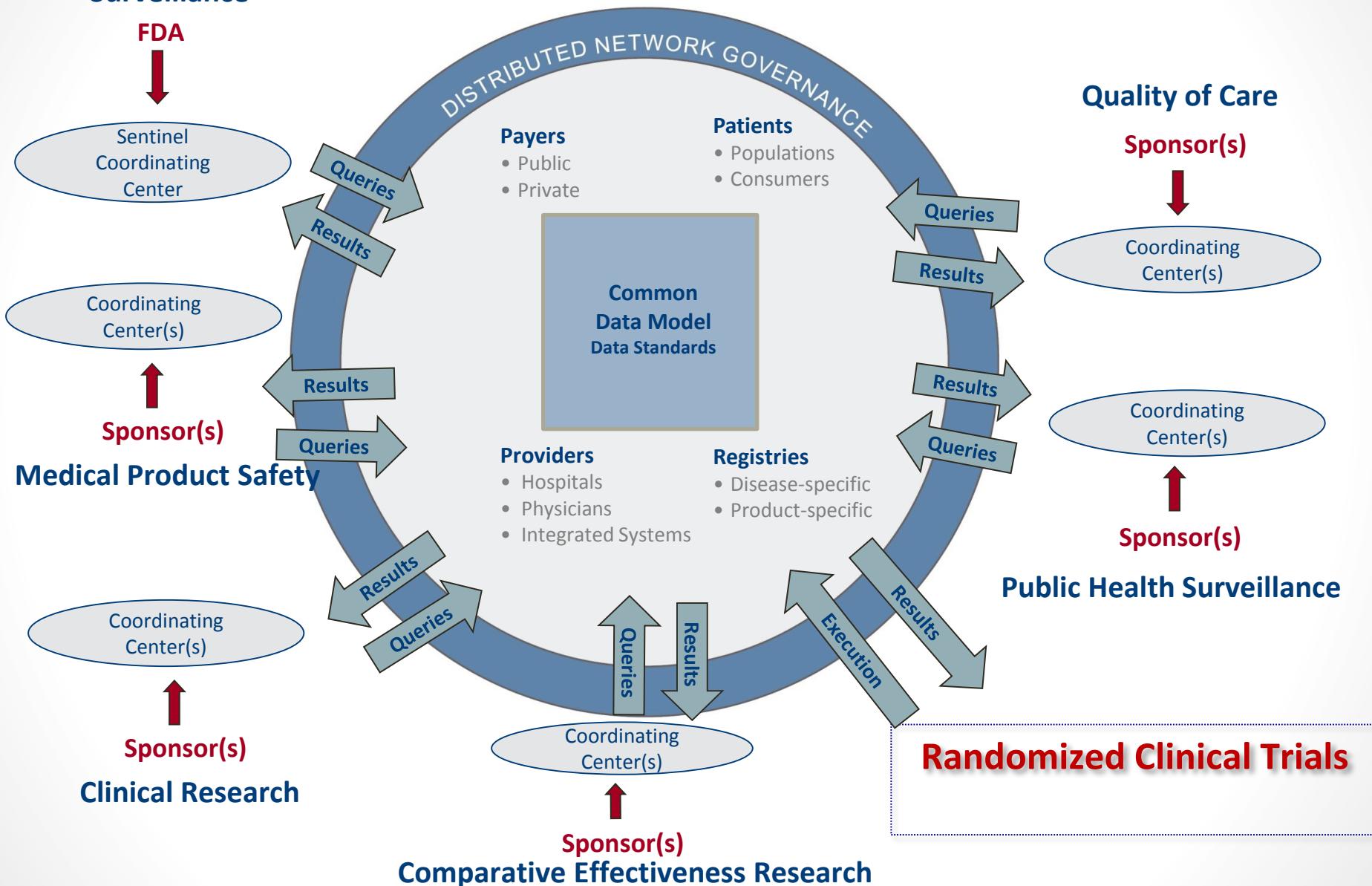
- Each organization can participate in multiple networks
- Each network controls its governance and coordination
- Networks share infrastructure, data curation, analytics, lessons, security, software development

Critical Partners in a National Infrastructure

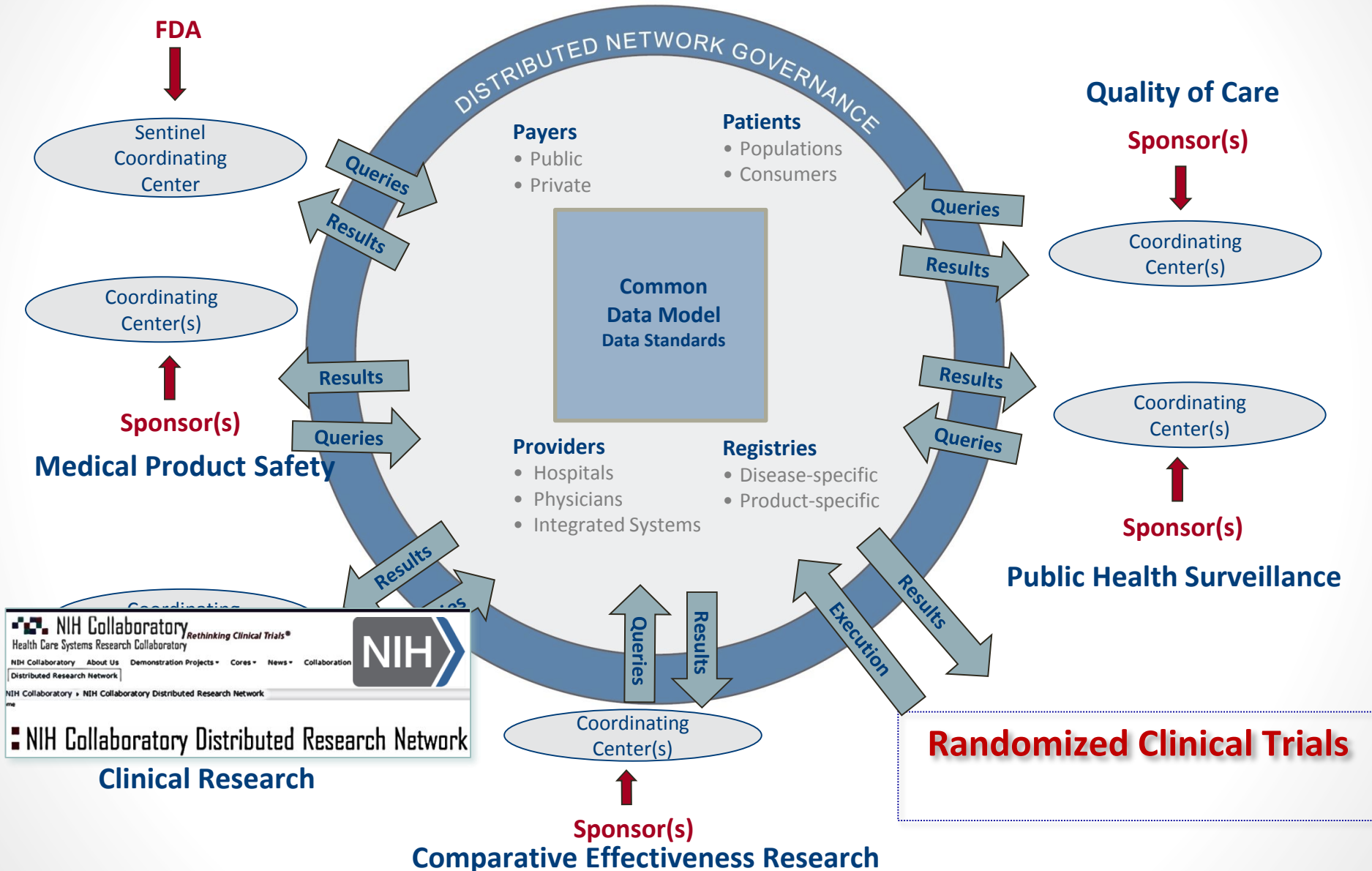


- Each organization can participate in multiple networks
- Each network controls its governance and coordination
- Networks share infrastructure, data curation, analytics, lessons, security, software development

Medical Product Safety National Evidence Generation Concept



Medical Product Safety National Evidence Generation Concept





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



Data Elements

- Available

- Ambulatory care diagnoses and procedures
- Outpatient pharmacy dispensing
- Laboratory testing and selected test results
- Inpatient diagnoses, treatments, and procedures itemized in hospital bill

- Not available


- Out-of-hospital death
- OTC medication
- Community-based immunizations

Pilot Test of the Collaboratory DRN

- Special NIH supplement in 2014 for pilot test
- Three pilot test queries developed by 3 NIH Institutes
- Pilot used publically-available Sentinel querying tools
- DRN Team and NIH staff (led by NHLBI & NCI) used queries as test cases for developing processes, and refining strategies to format queries
 - 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

Diabetes and Chelation Therapy

- Rationale: Assess recruitment feasibility of replicating the Trial to Assess Chelation Therapy (TACT)
- Goal: Characterize individuals with prevalent diabetes and prior AMI but no prior heart failure or chelation therapy
 - **Simple counts:** Counts and prevalence of chelation therapy and diabetes
 - **Complex counts:** First diagnosis of diabetes in people over 50 years of age in 2007 through 2014 with evidence of a prior AMI but no evidence of heart failure or chelation therapy
 - Any care setting
 - 365 day “look-back” window



Statin Users >75 years old and Cardiovascular Disease (CVD)

- Rationale: Characterize statin users over the age of 75 with regard to CVD and diabetes status
- **Complex counts:** All and long-term (≥ 180 days) prevalent and incident statin users
 - With no evidence of CVD
 - With and without a evidence of a diabetes diagnosis the day of or in the 90 days before first statin dispensing

Abnormal Cancer Screening Results and Follow-up

- Rationale: Characterize frequency of abnormal breast, colorectal, and cervical cancer screening test results and follow-up care
- **Background rates:** Incidence of cancer screenings and abnormal cancer screening results
 - 270 day “look-back” window to define new screen and new result
- **Abnormal screening results and follow-up:** For each cancer, count patients with a new abnormal finding, and among them, count how many had a follow-up test within 90 days
 - 183 day “look-back” window to define new abnormal result



Results: Cancer Screening and Abnormal Cancer Screen Result

- Breast cancer
 - **Screening (2007-2014)**
 - 6,719,382 eligible members (female, ages 40+, meets enrollment/incidence requirements, etc)
 - 3,750,337 new patients with a breast cancer screening
 - 8,809,583 new breast cancer screenings
 - **Abnormal Results (2007-2014)**
 - 6,898,880 eligible members (female, ages 40+, meets enrollment/incidence requirements, etc)
 - 1,075,964 patients with a new abnormal result
 - 1,418,562 new abnormal results



Results: Abnormal Cancer Screen Result and Follow-up

- Breast cancer continued...
 - **Follow-up after Abnormal Result (2013 only)**
 - 220,735 patients with a new abnormal result
 - 216,179 patients with a follow-up procedure/diagnosis
 - 97.9% follow-up within 90 days
 - 3.1 mean time to follow-up (days)



Results: Cancer Screening and Abnormal Cancer Screen Result

- Colorectal cancer
 - **Screening (2007-2014)**
 - 8,735,964 eligible members (ages 50+, meets enrollment/incidence requirements, etc)
 - 2,630,125 new patients with a colorectal cancer screening
 - 3,966,484 new colorectal cancer screenings
 - **Abnormal Results (2007-2014)**
 - 8,856,555 eligible members (ages 50+, meets enrollment/incidence requirements, etc)
 - 69,531 patients with a new abnormal result
 - 72,616 new abnormal results



Results: Abnormal Cancer Screen Result and Follow-up

- Colorectal cancer continued...
 - **Follow-up after Abnormal Result (2013 only)**
 - 12,121 patients with a new abnormal result
 - 8,545 patients with a follow-up procedure/diagnosis
 - 70.5% follow-up within 90 days
 - 32.0 mean time to follow-up (days)



Results: Cancer Screening and Abnormal Cancer Screen Result

- Cervical cancer
 - **Screening (2007-2014)**
 - 10,808,847 eligible members (female, ages 21+, meets enrollment/incidence requirements, etc)
 - 5,322,691 new patients with a cervical cancer screening
 - 10,703,839 new cervical cancer screenings
 - **Abnormal Results (2007-2014)**
 - 11,216,026 eligible members (female, ages 21+, meets enrollment/incidence requirements, etc)
 - 768,962 patients with a new abnormal result
 - 927,948 new abnormal results



Results: Abnormal Cancer Screen Result and Follow-up

- Cervical cancer continued...
 - **Follow-up after Abnormal Result (2013 only)**
 - 126,620 patients with a new abnormal result
 - 93,430 patients with a follow-up procedure/diagnosis
 - 73.8% follow-up within 90 days
 - 25.0 mean time to follow-up (days)



Summary of NIH Pilot Test

- Test cases assessed in three data organizations, representing $\sim 1/3$ of the total data
- Test cases informative of the necessary iterative process needed to refine queries
- Pilot informative of types of queries that are readily addressed vs. those that require a more iterative process over time to address
- Manual updated based on experience of the team with the test cases
- Revised processes and timelines for future test cases



Comparing Collaboratory DRN and PCORnet

- 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

<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 for "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 text below the heading states: "Millions of people. Strong collaborations. Privacy first." and "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 sidebar on the right contains three links: "DRN Governance Document, v1.0", "Distributed Research Network User's Guide, v5.0 (PDF)" (highlighted with a red box), 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 bottom of the screenshot shows the Windows taskbar with the Start button and several open applications: 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 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



Next Steps

- Exploring possibilities for testing additional NIH queries to refine process
- Considering pilot testing of external queries from the research community
- For additional information, please go to:
<https://www.nihcollaboratory.org/Pages/distributed-research-network.aspx#HowSubmit>