



NIH Collaboratory

Health Care Systems Research Collaboratory

# The NIH Collaboratory Distributed Research Network

Jeffrey Brown

Harvard Pilgrim Health Care Institute and

Harvard Medical School

June 5, 2015

# 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

<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 introductory text states: "Millions of people. Strong collaborations. Privacy first." and describes the network's purpose in enabling secure collaboration on electronic health data. A list of features is provided under the heading "What does the NIH Collaboratory Distributed Research Network do?". On the right side, there is a sidebar with three links: "DRN Governance Document, v1.0", "Distributed Research Network User's Guide, v5.0 (PDF)" (which is highlighted with a red rectangle), and "DRN Request Form (.docx)". At the bottom, a grey box contains contact information: "To learn more about the NIH Collaboratory Distributed Research Network" and the email address "support@popmednet.org". The Windows taskbar at the bottom shows the Start button and several open applications, including Outlook 2013 and the NIH Collaboratory Distributed Research Network.

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](mailto: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

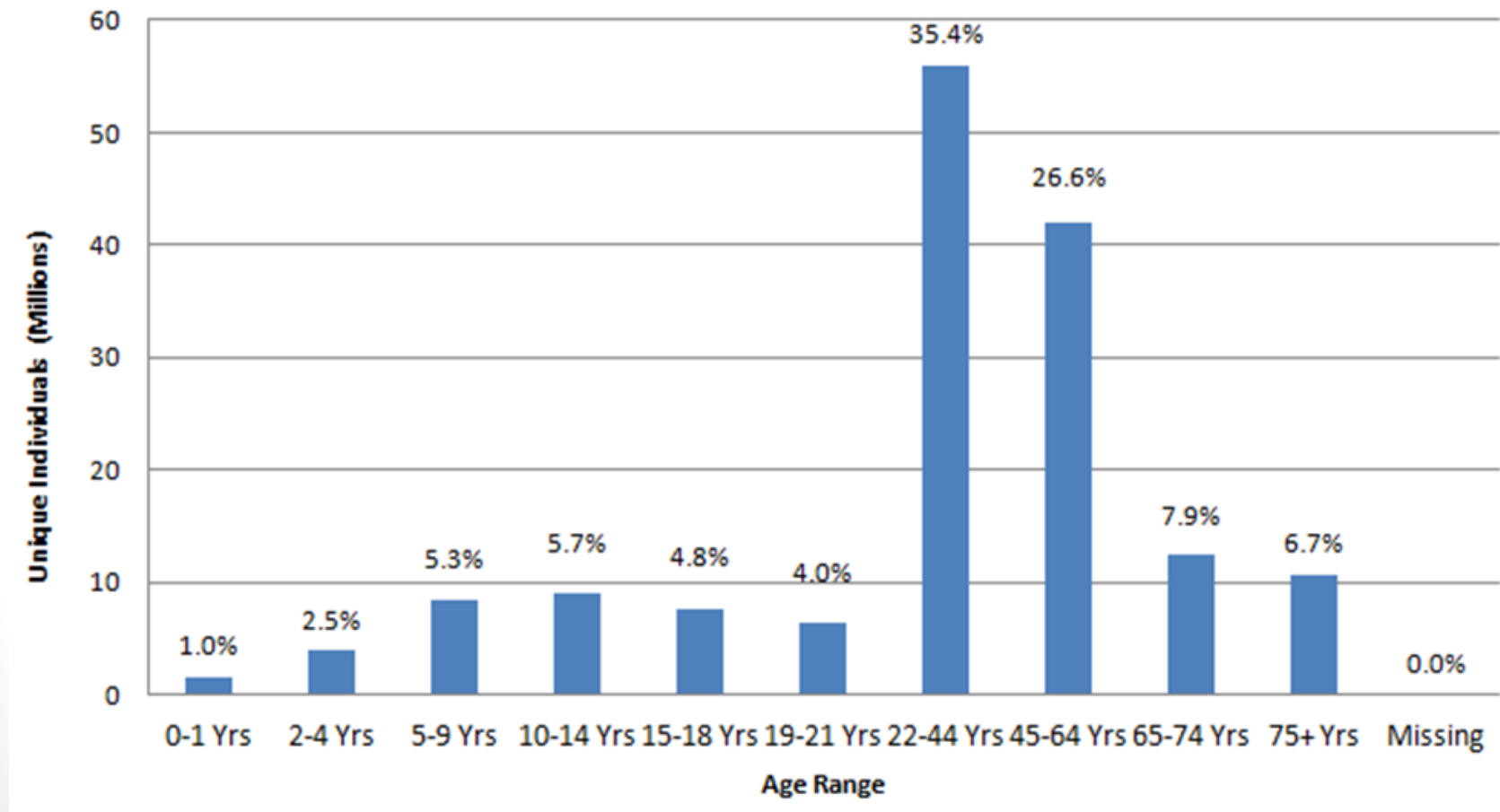
# Uses of the Distributed Network

- Provide information to support research planning
  - Background rates
  - Assess assumptions about relevant populations
  - Prioritize research domains
- Answer specific research questions
- Identify sites for participation in prospective interventional or observational studies

# Currently Available Data

- Research ready data sets representing >90% of the FDA Sentinel program
- > 300 million person-years of observation time and detailed information for billions of medical encounters and outpatient pharmacy dispensings

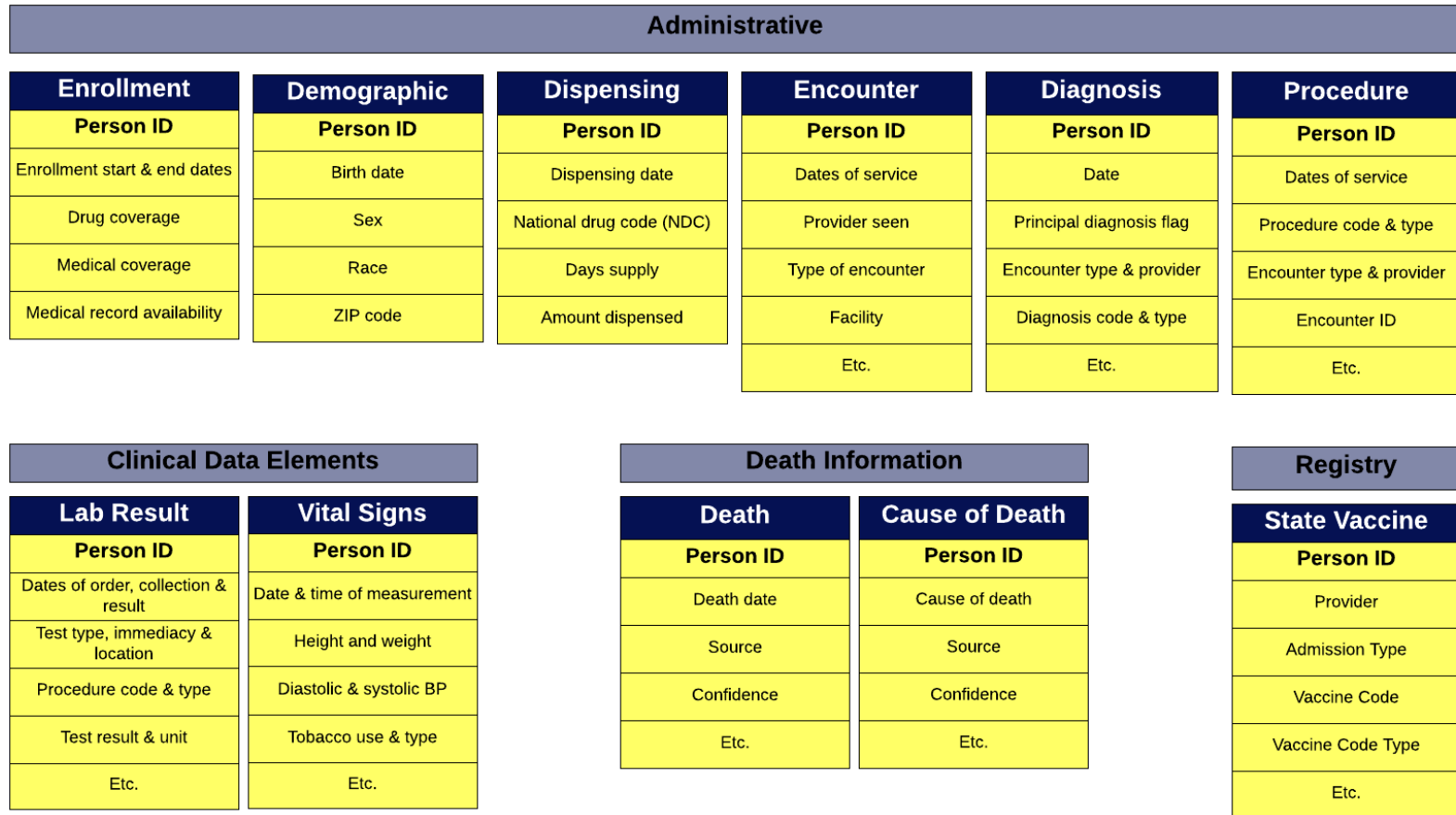
# Unique Individuals by Age Range



# Data Elements

- Captured
  - Ambulatory care diagnoses and procedures
  - Outpatient pharmacy dispensing
  - Laboratory testing and selected test results
  - Inpatient diagnoses, treatments and procedures itemized in hospital bill
- Not captured
  - Out of hospital death
  - Over-the-counter medication
  - Community-based immunizations

# Data Model



Some data partners do not create every table  
(e.g., vital signs are available for only a subset of individuals)



# The Easy – Hard Continuum of Questions

- Easy: Can be answered with existing programs
  - Counts, exposure-outcome relationships, confounder adjusted comparative cohort analyses

# The Easy – Hard Continuum of Questions

- Easy: Can be answered with existing programs
  - Counts, exposure-outcome relationships, confounder adjusted comparative cohort analyses
- Moderate: Can be answered with new programming
  - Data exists, is well characterized, and known to be reliable

# The Easy – Hard Continuum of Questions

- Easy: Can be answered with existing programs
  - Counts, exposure-outcome relationships, confounder adjusted comparative cohort analyses
- Moderate: Can be answered with new programming
  - Data exists, is well characterized, and known to be reliable
- Hard: Requires investigation or mapping of existing data
  - Data exists but completeness and quality must be determined

# The Easy – Hard Continuum of Questions

- Easy: Can be answered with existing programs
  - Counts, exposure-outcome relationships, confounder adjusted comparative cohort analyses
- Moderate: Can be answered with new programming
  - Data exists, is well characterized, and known to be reliable
- Hard: Requires investigation or mapping of existing data
  - Data exists but completeness and quality must be determined
- Harder: New data is needed
  - Birth registry, death registry, etc

# The Easy – Hard Continuum of Questions

- Easy: Can be answered with existing programs
  - Counts, exposure-outcome relationships, confounder adjusted comparative cohort analyses
- Moderate: Can be answered with new programming
  - Data exists, is well characterized, and known to be reliable
- Hard: Requires investigation or mapping of existing data
  - Data exists but completeness and quality must be determined
- Harder: New data is needed
  - Birth registry, death registry, etc
- Impossible: The data isn't reliably captured
  - Race, smoking status, over the counter medication use

# Where does the question fall on the continuum

- The DRN Coordinating Center helps requesters or their designees understand and use the network
- Assess fit between requests and the DRN's capabilities
- Suggest ways to maximize usefulness of the DRN data resources
- Facilitate engagement with data partners
- **Requesters do not have to be experts in observational research or use of health care data to initiate a request**

# Easy Example: Simple Counts

- Query goals
  - Counts of patients with Progressive Multifocal Leukoencephalopathy (PML)
- Analysis
  - Number of patients and prevalence rate of PML identified in inpatient setting
  - Counts provided per patient per year, age group, and sex

# Easy Example: Simple Counts

**Result:** In 2012, there were 87 individuals identified

## Prevalence of Progressive Multifocal Leukoencephalopathy in 2012

Age (years)	Males	Prevalence per 10,000	Females	Prevalence per 10,000
0-21	1	0.01	0	0
22-44	16	0.14	8	0.07
45-64	29	0.31	18	0.18
65+	6	0.16	9	0.20



# Easy Example:

## Cohort Identification and Descriptive Analysis

- Query goals
  - Patients continuously exposed to **bisphosphonates** for  $\geq 3$  years
  - Assess the risk of hip and other fractures
- Analysis
  - 2006 - 2013
  - Health plan members with medical and pharmacy coverage
  - **New** users of alendronate, risedronate, & ibandronate
  - Create treatment episodes based on repeated exposures
  - Identify fractures during or shortly after treatment
  - Sensitivity analyses examined different exposure, event, and episode definitions (n=78 analyses)

# Easy Example:

## Cohort Identification and Descriptive Analysis

### Results

- ~34,000 new users
- ~22,000 current alendronate users exposed for 3 - 5 years
- ~9,000 people enter this cohort each year

### Fractures in long term alendronate users\*

Fracture type	Exposed people	Person time (yrs)	Fractures	Rate / 10K yrs
Hip	34,428	138,386	725	52
Femoral fractures of interest	34,672	140,020	339	24

\* New users of alendronate, continuously exposed for at least 3 years

# Easy Example:

## Propensity score matched comparison

- Query goals
  - What is the comparative risk of angioedema among new users of ACE inhibitors vs. new users of beta-blockers?
- Analysis
  - Propensity score matched survival analysis
  - Performed via reusable modular program requiring only specification of input parameters

# Easy Example:

## Propensity score matched comparison

### Input parameters

- Population (age/sex/etc.), time period
- Exposures
- Outcomes
  - ICD-9-CM code 995.1 in any position during outpatient, inpatient, or emergency department encounter
  - Washout period (days before first dispensing): 183 days
- Inclusion criteria
- Exclusion criteria
- Covariates
- Propensity score matching options
  - Comorbidity, utilization, high dimensional propensity score
  - Matching ratio
  - Caliper size

# Angioedema: Table 1. Unmatched Cohort

Table 1. Cohort of New Initiators of ACE Inhibitors and Beta Blockers (Unmatched)

Characteristic	Primary Analysis				Covariate Balance	
	ACE Inhibitors		Beta Blockers		Absolute Difference	Standardized Difference
	N	%	N	%		
Patients	2,211,215	100%	1,673,682	100%	0.0	-
Events while on therapy	5,158	0.2%	1,292	0.1%	0.1	0.0
Person-time at risk (days)	186.9	266.6	149.2	235.1	37.7	0.2
<b>Patient Characteristics</b>						
Gender (F)	997,962	45.10%	946,344	56.50%	-11.4	-0.2
Mean age (std dev)	54.6	12.7	53.7	15.6	0.9	0.1
<b>Recorded History of:</b>						
Allergic reactions	207,344	9.4%	190,387	11.4%	-2.0	-0.1
Diabetes	471,661	21.3%	173,083	10.3%	11.0	0.3
Heart failure	41,060	1.9%	74,897	4.5%	-2.6	-0.1
Ischemic heart diseases	109,948	5.0%	224,681	13.4%	-8.4	-0.3
NSAID use	318,298	14.4%	250,697	15.0%	-0.6	0.0
<b>Health Service Utilization Intensity:</b>						
	Mean	Std Dev	Mean	Std Dev		
Number of generics	3.4	3.5	4.1	4.0	-0.7	-0.2
Number of filled prescriptions	7.5	9.6	8.9	10.8	-1.4	-0.1
Number of inpatient hospital encounters (IP)	0.1	0.4	0.2	0.6	-0.1	-0.3
Number of non-acute institutional encounters (IS)	0.0	0.6	0.1	0.9	-0.1	-0.1
Number of emergency room encounters (ED)	0.2	0.7	0.4	1.0	-0.2	-0.2
Number of ambulatory encounters (AV)	4.8	6.3	6.9	8.4	-2.1	-0.3
Number of other ambulatory encounters (OA)	1.1	2.6	1.5	3.6	-0.4	-0.1

[www.mini-sentinel.org/work\\_products/Statistical\\_Methods/Mini-Sentinel\\_Methods\\_Known-Positives-ACEI-Angioedema.pdf](http://www.mini-sentinel.org/work_products/Statistical_Methods/Mini-Sentinel_Methods_Known-Positives-ACEI-Angioedema.pdf)

# Angioedema: Table 1. Unmatched Cohort

Table 1. Cohort of New Initiators of ACE Inhibitors and Beta Blockers (Unmatched)

Characteristics	Primary Analysis				Covariate Balance	
	ACE Inhibitors	Beta Blockers			Absolute Difference	Standardized Difference
		%	N	%		
Patients	2,211,215	100%	1,673,682	100%	0.0	-
Events while on therapy	5,158	0.2%	1,292	0.1%	0.1	0.0
Person-time at risk (days)	186.9	266.6	149.2	235.1	37.7	0.2
<b>Patient Characteristics</b>						
Gender (F)	997,962	45.10%	946,344	56.50%	-11.4	-0.2
Mean age (std dev)	54.6	12.7	53.7	15.6	0.9	0.1
<b>Recorded History of:</b>						
Allergic reactions	207,344	9.4%	190,387	11.4%	-2.0	-0.1
Diabetes	471,661	21.3%	173,083	10.3%	11.0	0.3
Heart failure	41,060	1.9%	74,897	4.5%	-2.6	-0.1
Ischemic heart diseases	109,948	5.0%	224,681	13.4%	-8.4	-0.3
NSAID use	318,298	14.4%	250,697	15.0%	-0.6	0.0
<b>Health Service Utilization Intensity:</b>						
	Mean	Std Dev	Mean	Std Dev		
Number of generics	3.4	3.5	4.1	4.0	-0.7	-0.2
Number of filled prescriptions	7.5	9.6	8.9	10.8	-1.4	-0.1
Number of inpatient hospital encounters (IP)	0.1	0.4	0.2	0.6	-0.1	-0.3
Number of non-acute institutional encounters (IS)	0.0	0.6	0.1	0.9	-0.1	-0.1
Number of emergency room encounters (ED)	0.2	0.7	0.4	1.0	-0.2	-0.2
Number of ambulatory encounters (AV)	4.8	6.3	6.9	8.4	-2.1	-0.3
Number of other ambulatory encounters (OA)	1.1	2.6	1.5	3.6	-0.4	-0.1

[www.mini-sentinel.org/work\\_products/Statistical\\_Methods/Mini-Sentinel\\_Methods\\_Known-Positives-ACEI-Angioedema.pdf](http://www.mini-sentinel.org/work_products/Statistical_Methods/Mini-Sentinel_Methods_Known-Positives-ACEI-Angioedema.pdf)

# Angioedema: Table 1. Unmatched Cohort

Table 1. Cohort of New Initiators of ACE Inhibitors and Beta Blockers (Unmatched)

Characteristic	Primary Analysis				Covariate Balance	
	ACE Inhibitors		Beta Blockers		Absolute Difference	Standardized Difference
		%	N	%		
Patients	2,211,215	100%	1,673,682	100%	0.0	-
Events while on therapy	5,158	0.2%	1,292	0.1%	0.1	0.0
Person-time at risk (days)	186.9	266.6	149.2	235.1	37.7	0.2
<b>Patient Characteristics</b>						
Gender (F)	997,962	45.10%	946,344	56.50%	-11.4	-0.2
Mean age (std dev)	54.6	12.7	53.7	15.6	0.9	0.1
<b>Recorded History of:</b>						
Allergic reactions	207,344	9.4%	190,387	11.4%	-2.0	-0.1
Diabetes	471,661	21.3%	173,083	10.3%	11.0	0.3
Heart failure	41,060	1.9%	74,897	4.5%	-2.6	-0.1
Ischemic heart diseases	109,948	5.0%	224,681	13.4%	-8.4	-0.3
NSAID use	318,298	14.4%	250,697	15.0%	-0.6	0.0
<b>Health Service Utilization Intensity:</b>						
	Mean	Std Dev				
Number of generics	3.4	3.5				
Number of filled prescriptions	7.5	9.6				
Number of inpatient hospital encounters (IP)	0.1	0.4				
Number of non-acute institutional encounters (IS)	0.0	0.6	0.1	0.9	-0.1	-0.1
Number of emergency room encounters (ED)	0.2	0.7	0.4	1.0	-0.2	-0.2
Number of ambulatory encounters (AV)	4.8	6.3	6.9	8.4	-2.1	-0.3
Number of other ambulatory encounters (OA)	1.1	2.6	1.5	3.6	-0.4	-0.1

3.9 million new users

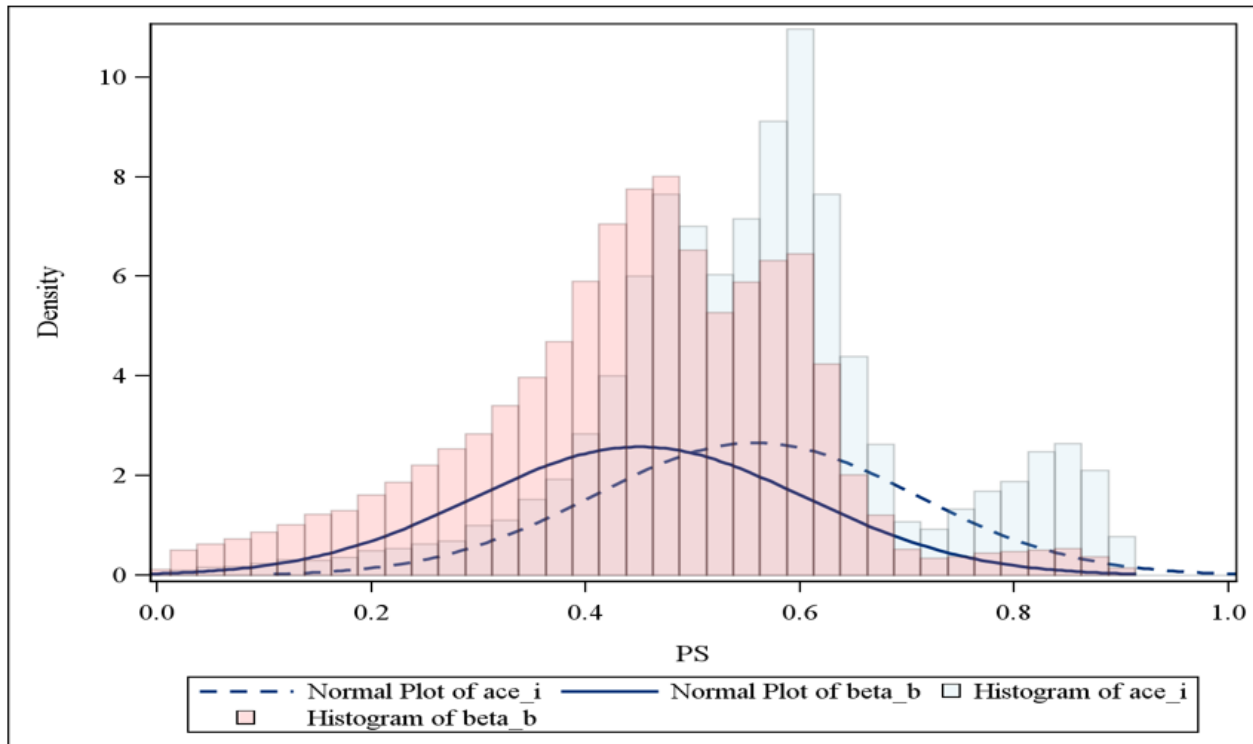
Diabetes 21% vs 10%  
Heart failure 2% vs 4%  
Ischemic heart disease 5% vs 13%

[www.mini-sentinel.org/work\\_products/Statistical\\_Methods/Mini-Sentinel\\_Methods\\_Known-Positives-ACEI-Angioedema.pdf](http://www.mini-sentinel.org/work_products/Statistical_Methods/Mini-Sentinel_Methods_Known-Positives-ACEI-Angioedema.pdf)

# Propensity Scores Before Match

## Histograms of PS distribution by DP (masked)

Histogram of Predefined PS, Unmatched Cohort C-Stat for Predefined: 0.695



[www.mini-sentinel.org/work\\_products/Statistical\\_Methods/Mini-Sentinel\\_Methods\\_Known-Positives-ACEI-Angioedema.pdf](http://www.mini-sentinel.org/work_products/Statistical_Methods/Mini-Sentinel_Methods_Known-Positives-ACEI-Angioedema.pdf)



# Angioedema: Table 2. Matched Cohort

Table 2. Cohort of New Initiators of ACE Inhibitors and Beta Blockers (Matched Predefined PS, Caliper = .025)

Characteristics	Primary Analysis				Covariate Balance	
	ACE Inhibitors		Beta Blockers		Absolute Difference	Standardized Difference
<b>2.6 million new users</b>	N	%	N	%		
Patients	1,309,104	59.2%	1,309,104	78.2%	0.0	-0.4
Events while on therapy	3,311	0.3%	988	0.1%	0.2	0.0
Person-time at risk (days)	183.8	263.7	151.8	238.9	31.9	0.1
<b>Patient Characteristics</b>						
Gender (F)	723,955	55.3%	689,617	52.7%	2.6	0.1
Mean age (std dev)	54.1	13.1	54.4	14.9	-0.3	0.0
<b>Recorded History of:</b>						
Allergic reactions	137,920	10.5%	134,933	10.3%	0.2	0.0
Diabetes	150,036	11.5%	150,551	11.5%	0.0	0.0
Heart failure	35,302	2.7%	38,966	3.0%	-0.3	0.0
Ischemic heart diseases	102,200	7.8%	106,786	8.2%	-0.4	0.0
NSAID use	191,798	14.7%	189,612	14.5%	0.2	0.0
<b>Health Service Utilization Intensity:</b>						
	Mean	Std				
Number of generics	3.7	3.7%				
Number of filled prescriptions	8.1	10.2%				
Number of inpatient hospital encounters (IP)	0.1	0.5%				
Number of non-acute institutional encounters (IS)	0.1	0.7%	0.1	0.7%	0.0	0.0
Number of emergency room encounters (ED)	0.3	0.8%	0.3	0.8%	0.0	0.0
Number of ambulatory encounters (AV)	5.6	7.3%	5.6	6.6%	0.0	0.0
Number of other ambulatory encounters (OA)	1.2	2.9%	1.3	3.0%	0.0	0.0

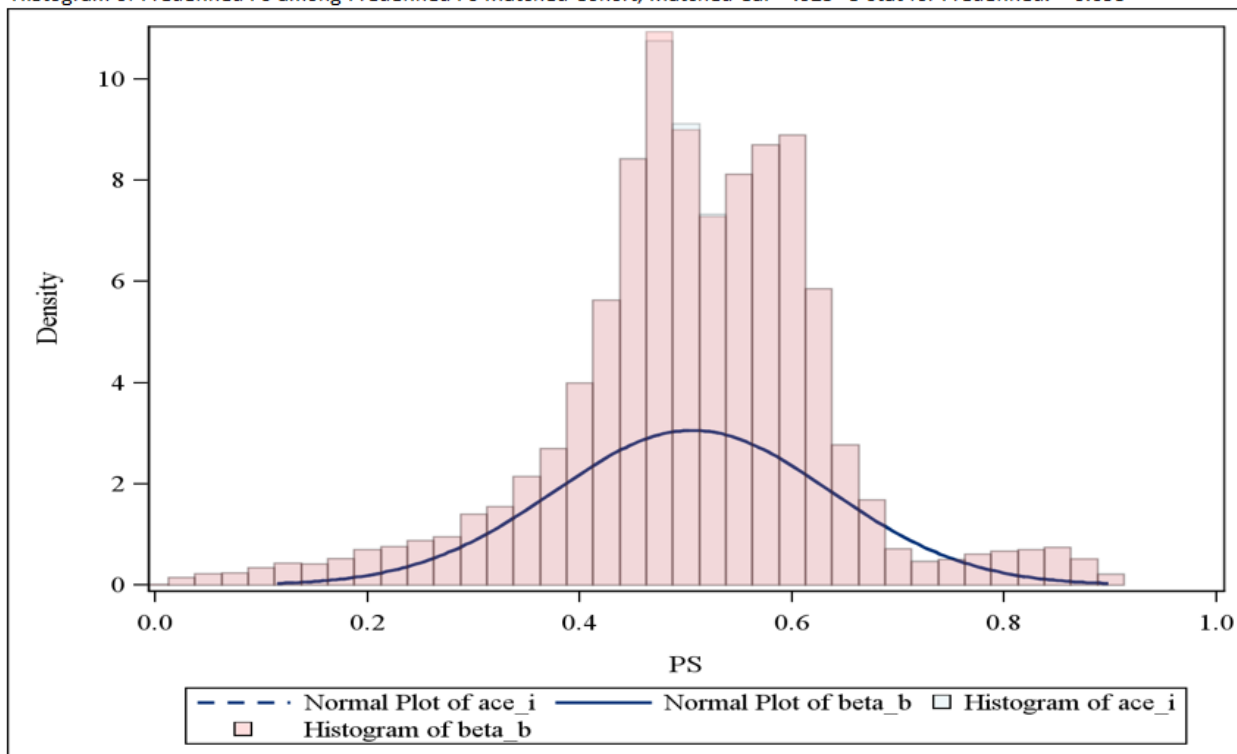
**Diabetes** 10% vs 10%  
**Heart failure** 3% vs 3%  
**Ischemic heart disease** 8% vs 8%

[www.mini-sentinel.org/work\\_products/Statistical\\_Methods/Mini-Sentinel\\_Methods\\_Known-Positives-ACEI-Angioedema.pdf](http://www.mini-sentinel.org/work_products/Statistical_Methods/Mini-Sentinel_Methods_Known-Positives-ACEI-Angioedema.pdf)

# Propensity Scores After Match

## Histograms of PS distribution by DP (masked)

Histogram of Predefined PS among Predefined PS Matched Cohort, Matched Cal = .025 C-Stat for Predefined: 0.695



[www.mini-sentinel.org/work\\_products/Statistical\\_Methods/Mini-Sentinel\\_Methods\\_Known-Positives-ACEI-Angioedema.pdf](http://www.mini-sentinel.org/work_products/Statistical_Methods/Mini-Sentinel_Methods_Known-Positives-ACEI-Angioedema.pdf)

# Angioedema: Table 3. Results

Table 3: Sequential Estimates for Angioedema Events by Analysis Type, and Drug Pair					
Exposure Definition	Monitoring Period	New Users	Person Years at Risk	Average Person Years at Risk	Number of Events
Unmatched Analysis (Site-adjusted only)					
ACE Inhibitors	1	2,211,215	1,131,526	0.51	5,158
Beta Blockers		1,673,682	683,614	0.41	1,292
1:1 Matched Analysis; Caliper=0.025					
ACE Inhibitors	1	1,309,104	658,700	0.50	3,311
Beta Blockers		1,309,104	544,285	0.42	988

Incidence Rate per 1000 Person Years	Risk per 1000 New Users	Difference per 1000 Person Years	Difference in Risk per 1000 New Users	Hazard Ratio (95% CI)	Wald P-Value
4.558	2.33	2.67	1.56	2.55 ( 2.40, 2.71)	<.0001
1.890	0.77				
5.027	2.53	3.21	1.77	3.14 ( 2.86, 3.44)	<.0001
1.815	0.75				

[www.mini-sentinel.org/work\\_products/Statistical\\_Methods/Mini-Sentinel\\_Methods\\_Known-Positives-ACEI-Angioedema.pdf](http://www.mini-sentinel.org/work_products/Statistical_Methods/Mini-Sentinel_Methods_Known-Positives-ACEI-Angioedema.pdf)

# Angioedema: Table 3. Results

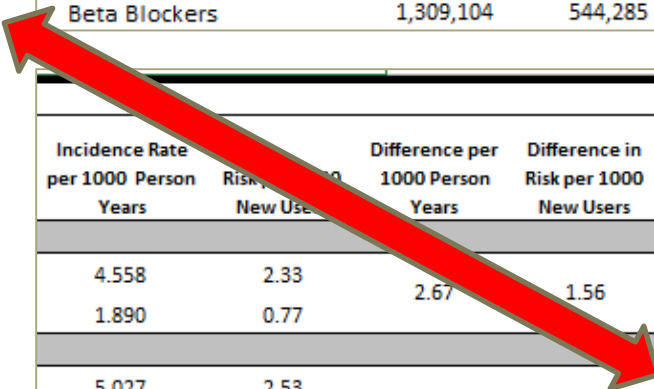
ACEI vs  $\beta$ -blocker 1:1  
matched analysis:

- **HR = 3.1**  
(95% CI, 2.9-3.4)

Exposure Definition	Monitoring Period	New Users	Person Years at Risk	Average Person Years at Risk	Number of Events
Unmatched Analysis (Site-adjusted only)					
ACE Inhibitors	1	2,211,215	1,131,526	0.51	5,158
Beta Blockers		1,673,682	683,614	0.41	1,292
1:1 Matched Analysis; Caliper=0.025					
ACE Inhibitors	1	1,309,104	658,700	0.50	3,311
Beta Blockers		1,309,104	544,285	0.42	988

Incidence Rate per 1000 Person Years	Risk per 1000 New Users	Difference per 1000 Person Years	Difference in Risk per 1000 New Users	Hazard Ratio (95% CI)	Wald P-Value
4.558	2.33	2.67	1.56	2.55 ( 2.40, 2.71)	<.0001
1.890	0.77				
5.027	2.53	3.21	1.77	3.14 ( 2.86, 3.44)	<.0001
1.815	0.75				



[www.mini-sentinel.org/work\\_products/Statistical\\_Methods/Mini-Sentinel\\_Methods\\_Known-Positives-ACEI-Angioedema.pdf](http://www.mini-sentinel.org/work_products/Statistical_Methods/Mini-Sentinel_Methods_Known-Positives-ACEI-Angioedema.pdf)

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

- Plan to replicate the TACT trial – EDTA chelation to prevent coronary heart disease – focusing on diabetic patients
- **Inclusion criteria**
  - > 50 years old
  - Confirmed diagnosis of diabetes on medical therapy (insulin or oral)
  - Previous myocardial infarction

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

- Plan to replicate the TACT trial – EDTA chelation to prevent coronary heart disease – focusing on diabetic patients
- **Inclusion criteria**
  - > 50 years old
  - Confirmed diagnosis of diabetes on medical therapy (insulin or oral)
  - Previous myocardial infarction

EASY: All inclusion criteria are available for querying using existing cohort identification programs

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

### Exclusion criteria

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

### Exclusion criteria

- Creatinine > 2.0 mg/dl



# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

### Exclusion criteria

- Creatinine > 2.0 mg/dl
  - [EASY](#): Available for a subset; >7million results available

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

### Exclusion criteria

- Creatinine > 2.0 mg/dl
  - [EASY](#): Available for a subset; >7million results available
- Cigarette smoking within 3 months

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

### Exclusion criteria

- Creatinine > 2.0 mg/dl
  - EASY: Available for a subset; >7million results available
- Cigarette smoking within 3 months
  - IMPOSSIBLE: Smoking status not recorded in claims and unreliable in EHRs

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

### Exclusion criteria

- Creatinine > 2.0 mg/dl
  - EASY: Available for a subset; >7million results available
- Cigarette smoking within 3 months
  - IMPOSSIBLE: Smoking status not recorded in claims and unreliable in EHRs
- Heart failure or heart failure hospitalization

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

### Exclusion criteria

- Creatinine > 2.0 mg/dl
  - EASY: Available for a subset; >7million results available
- Cigarette smoking within 3 months
  - IMPOSSIBLE: Smoking status not recorded in claims and unreliable in EHRs
- Heart failure or heart failure hospitalization
  - EASY: Available

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

### Exclusion criteria

- Creatinine > 2.0 mg/dl
  - [EASY](#): Available for a subset; >7million results available
- Cigarette smoking within 3 months
  - [IMPOSSIBLE](#): Smoking status not recorded in claims and unreliable in EHRs
- Heart failure or heart failure hospitalization
  - [EASY](#): Available
- No chelation therapy in prior 5 years

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

### Exclusion criteria

- Creatinine > 2.0 mg/dl
  - EASY: Available for a subset; >7million results available
- Cigarette smoking within 3 months
  - IMPOSSIBLE: Smoking status not recorded in claims and unreliable in EHRs
- Heart failure or heart failure hospitalization
  - EASY: Available
- No chelation therapy in prior 5 years
  - Probably EASY: Need to assess data capture reliability and payment policies

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

- Question: What are the demographic characteristics of patients that might be eligible – race, gender, age? What about comorbidities?



# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

- Question: What are the demographic characteristics of patients that might be eligible – race, gender, age? What about comorbidities?
  - EASY: Age, sex, and comorbidities can be defined and presented

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

- Question: What are the demographic characteristics of patients that might be eligible – race, gender, age? What about comorbidities?
  - EASY: Age, sex, and comorbidities can be defined and presented
  - IMPOSSIBLE: Race is recorded for a subset of patients

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

- Question: What can you tell us about where patients who meet these criteria receive most of their care – primary care offices, cardiology offices, endocrinology clinics? Does this vary in urban, suburban, more rural communities?

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

- Question: What can you tell us about where patients who meet these criteria receive most of their care – primary care offices, cardiology offices, endocrinology clinics? Does this vary in urban, suburban, more rural communities?
  - HARD: Facility and provider codes are available; new programming and discussion with data partners would be required

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

- What can you tell us about the uncertainties in these estimates?

# Example Request Assessment

## Trial to Assess Chelation Therapy (TACT) Replication

- What can you tell us about the uncertainties in these estimates?
  - Suggest using sensitivity analyses to assess importance of each definition

# Example Request Assessment

## Follow Up of Abnormal Cancer Screening Tests

Request: Characterize rate of follow-up of abnormal cancer screening tests, including mammography, fecal immunochemical (FIT), or Pap tests within a managed care population

# Example Request Assessment

## Follow Up of Abnormal Cancer Screening Tests

- Identification of benefit design – to define “managed care” – is possible but complex



# Example Request Assessment

## Follow Up of Abnormal Cancer Screening Tests

- Identification of benefit design – to define “managed care” – is possible but complex
  - Assessment of complexity and validity over time is needed
  - Definition of “managed care”

# Example Request Assessment

## Follow Up of Abnormal Cancer Screening Tests

1. How many are screened for each cancer?
2. How many have abnormal screening test results?
3. How many abnormal results appear to have no further testing?
  - a. For mammography – no additional mammography, ultrasound, MRI or biopsy with 90 days
  - b. For FIT – no colonoscopy within 90 days
  - c. For PAP – no repeat PAP that is normal, or no colposcopy within 90 days
4. Is there other evidence of evaluation of the abnormality?

# Example Request Assessment

## Follow Up of Abnormal Cancer Screening Tests

1. How many are screened for each cancer?
2. How many have abnormal screening test results?
3. How many abnormal results appear to have no further testing?
  - a. For mammography – no additional mammography, ultrasound, MRI or biopsy with 90 days
  - b. For FIT – no colonoscopy within 90 days
  - c. For PAP – no repeat PAP that is normal, or no colposcopy within 90 days
4. Is there other evidence of evaluation of the abnormality?

EASY: Questions 1-4 can be answered using existing data and programs

# Example Request Assessment

## Follow Up of Abnormal Cancer Screening Tests

5. Does the rate of follow up of abnormal test results vary across practices?

# Example Request Assessment

## Follow Up of Abnormal Cancer Screening Tests

5. Does the rate of follow up of abnormal test results vary across practices?

HARD: Facility and provider codes are available; new programming and discussion with data partners would be required

# Example Request Assessment

## Follow Up of Abnormal Cancer Screening Tests

5. Does the rate of follow up of abnormal test results vary across practices?

HARD: Facility and provider codes are available; new programming and discussion with data partners would be required

What are the race and age breakdowns of patients?

# Example Request Assessment

## Follow Up of Abnormal Cancer Screening Tests

5. Does the rate of follow up of abnormal test results vary across practices?

HARD: Facility and provider codes are available; new programming and discussion with data partners would be required

What are the race and age breakdowns of patients?

- EASY: Age distribution

# Example Request Assessment

## Follow Up of Abnormal Cancer Screening Tests

5. Does the rate of follow up of abnormal test results vary across practices?

HARD: Facility and provider codes are available; new programming and discussion with data partners would be required

What are the race and age breakdowns of patients?

- EASY: Age distribution
- IMPOSSIBLE: Race



# How to Use the NIH Collaboratory Distributed Research Network

- Data Partners participate on a project-by-project-basis
- Submit requests using the [NIH Collaboratory DRN request form](#)
- The DRN Coordinating Center reviews each request to assess appropriateness and level of effort
- Costs: on a case-by-case basis

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

The screenshot shows a Firefox browser window with the address bar displaying <https://www.nihcollaboratory.org/Pages/distributed-research-network.aspx>. The website header features the NIH Collaboratory logo with the tagline "Rethinking Clinical Trials®" and the subtitle "Health Care Systems Research Collaboratory". A search bar is located on the right. The navigation menu includes links for "NIH Collaboratory", "About Us", "Demonstration Projects", "Cores", "News", "Collaboration Spaces", "Knowledge Repository", "The Living Textbook", and "Grand Rounds". The main content area is titled "NIH Collaboratory Distributed Research Network" and features the tagline "Millions of people. Strong collaborations. Privacy first." Below this, a paragraph 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 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 provides contact information: "To learn more about the NIH Collaboratory Distributed Research Network" and the email address "support@popmednet.org". The bottom of the browser window shows the Windows taskbar with the Start button and several open applications, including Outlook 2013 and the NIH Collaboratory Distributed Research Network.

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](mailto: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

Thank you!