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.
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@pcpmednet.org

DRN Governance Document, v1.0
Distributed Research Network User’s Guide, v5.0 (PDF)
DRN Request Form (.docx)
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
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

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Males</th>
<th>Prevalence per 10,000</th>
<th>Females</th>
<th>Prevalence per 10,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-21</td>
<td>1</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22-44</td>
<td>16</td>
<td>0.14</td>
<td>8</td>
<td>0.07</td>
</tr>
<tr>
<td>45-64</td>
<td>29</td>
<td>0.31</td>
<td>18</td>
<td>0.18</td>
</tr>
<tr>
<td>65+</td>
<td>6</td>
<td>0.16</td>
<td>9</td>
<td>0.20</td>
</tr>
</tbody>
</table>
Easy Example: Cohort Identification and Descriptive Analysis

• Query goals
  • Patients continuously exposed to **bisphosphonates** for ≥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

<table>
<thead>
<tr>
<th>Fracture type</th>
<th>Exposed people</th>
<th>Person time (yrs)</th>
<th>Fractures</th>
<th>Rate / 10K yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hip</td>
<td>34,428</td>
<td>138,386</td>
<td>725</td>
<td>52</td>
</tr>
<tr>
<td>Femoral fractures of interest</td>
<td>34,672</td>
<td>140,020</td>
<td>339</td>
<td>24</td>
</tr>
</tbody>
</table>

* 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)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ACE Inhibitors</th>
<th>Beta Blockers</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Patients</td>
<td>2,211,215</td>
<td>100%</td>
<td>1,673,682</td>
<td>100%</td>
</tr>
<tr>
<td>Events while on therapy</td>
<td>5,158</td>
<td>0.2%</td>
<td>1,292</td>
<td>0.1%</td>
</tr>
<tr>
<td>Person-time at risk (days)</td>
<td>186.9</td>
<td>266.6</td>
<td>149.2</td>
<td>235.1</td>
</tr>
</tbody>
</table>

### Patient Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ACE Inhibitors</th>
<th>Beta Blockers</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (F)</td>
<td>997,952</td>
<td>45.10%</td>
<td>946,344</td>
<td>56.50%</td>
</tr>
<tr>
<td>Mean age (std dev)</td>
<td>54.6</td>
<td>12.7</td>
<td>53.7</td>
<td>15.6</td>
</tr>
</tbody>
</table>

### Recorded History:

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ACE Inhibitors</th>
<th>Beta Blockers</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allergic reactions</td>
<td>207,344</td>
<td>9.4%</td>
<td>190,387</td>
<td>11.4%</td>
</tr>
<tr>
<td>Diabetes</td>
<td>471,661</td>
<td>21.3%</td>
<td>173,083</td>
<td>10.3%</td>
</tr>
<tr>
<td>Heart failure</td>
<td>41,060</td>
<td>1.9%</td>
<td>74,897</td>
<td>4.8%</td>
</tr>
<tr>
<td>Ischemic heart diseases</td>
<td>109,948</td>
<td>5.0%</td>
<td>224,681</td>
<td>13.4%</td>
</tr>
<tr>
<td>NSAID use</td>
<td>318,238</td>
<td>14.4%</td>
<td>250,857</td>
<td>15.0%</td>
</tr>
</tbody>
</table>

### Health Service Utilization Intensity:

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ACE Inhibitors</th>
<th>Beta Blockers</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of generics</td>
<td>5.4</td>
<td>3.5</td>
<td>4.1</td>
<td>4.0</td>
</tr>
<tr>
<td>Number of filled prescriptions</td>
<td>7.5</td>
<td>9.6</td>
<td>8.9</td>
<td>10.8</td>
</tr>
<tr>
<td>Number of inpatient hospital encounters</td>
<td>0.1</td>
<td>0.4</td>
<td>0.2</td>
<td>0.6</td>
</tr>
<tr>
<td>Number of non-acute institutional encounters</td>
<td>0.0</td>
<td>0.6</td>
<td>0.1</td>
<td>0.9</td>
</tr>
<tr>
<td>Number of emergency room encounters (ED)</td>
<td>0.2</td>
<td>0.7</td>
<td>0.4</td>
<td>1.0</td>
</tr>
<tr>
<td>Number of ambulatory encounters (AV)</td>
<td>4.8</td>
<td>6.3</td>
<td>6.9</td>
<td>8.4</td>
</tr>
<tr>
<td>Number of other ambulatory encounters (OA)</td>
<td>1.1</td>
<td>2.6</td>
<td>1.5</td>
<td>3.6</td>
</tr>
</tbody>
</table>

Angioedema: Table 1. Unmatched Cohort

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ACE Inhibitors</th>
<th>Beta Blockers</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients</td>
<td>2,211,215</td>
<td>1,673,632</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Patients on therapy</td>
<td>5,158</td>
<td>1,291</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Person-time at risk (days)</td>
<td>186.9</td>
<td>149.2</td>
<td>23.1</td>
<td>37.7</td>
</tr>
</tbody>
</table>

3.9 million new users

<table>
<thead>
<tr>
<th>Patient Characteristics</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (F)</td>
<td>937,952</td>
<td>45.1%</td>
<td>946,344</td>
<td>56.5%</td>
<td>-11.4</td>
<td>-0.2</td>
</tr>
<tr>
<td>Mean age (std dev)</td>
<td>54.6</td>
<td>12.7</td>
<td>53.7</td>
<td>15.6</td>
<td>0.9</td>
<td>0.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Recorded History of</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Allergic reactions</td>
<td>207,344</td>
<td>9.4%</td>
<td>190,387</td>
<td>11.4%</td>
<td>-2.0</td>
<td>-0.1</td>
</tr>
<tr>
<td>Diabetes</td>
<td>471,661</td>
<td>21.3%</td>
<td>175,083</td>
<td>10.3%</td>
<td>11.0</td>
<td>0.3</td>
</tr>
<tr>
<td>Heart failure</td>
<td>41,960</td>
<td>1.9%</td>
<td>74,897</td>
<td>4.5%</td>
<td>-2.6</td>
<td>-0.1</td>
</tr>
<tr>
<td>Ischemic heart diseases</td>
<td>109,948</td>
<td>5.0%</td>
<td>224,681</td>
<td>13.4%</td>
<td>6.4</td>
<td>0.3</td>
</tr>
<tr>
<td>NSAID use</td>
<td>318,238</td>
<td>14.4%</td>
<td>250,857</td>
<td>15.0%</td>
<td>-0.6</td>
<td>0.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Health Service Utilization Intensity:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of generics</td>
<td>5.4</td>
<td>3.5%</td>
<td>4.1</td>
<td>4.0%</td>
<td>-0.7</td>
<td>-0.2</td>
</tr>
<tr>
<td>Number of filled prescriptions</td>
<td>7.5</td>
<td>9.6%</td>
<td>8.9</td>
<td>10.8%</td>
<td>-1.4</td>
<td>-0.1</td>
</tr>
<tr>
<td>Number of inpatient hospital encounters (IP)</td>
<td>0.1</td>
<td>0.4%</td>
<td>0.2</td>
<td>0.5%</td>
<td>-0.1</td>
<td>-0.3</td>
</tr>
<tr>
<td>Number of non-acute institutional encounters (IS)</td>
<td>0.0</td>
<td>0.6%</td>
<td>0.1</td>
<td>0.9%</td>
<td>-0.1</td>
<td>-0.1</td>
</tr>
<tr>
<td>Number of emergency room encounters (ED)</td>
<td>0.2</td>
<td>0.7%</td>
<td>0.4</td>
<td>1.0%</td>
<td>-0.2</td>
<td>-0.2</td>
</tr>
<tr>
<td>Number of ambulatory encounters (AV)</td>
<td>4.8</td>
<td>6.3%</td>
<td>6.9</td>
<td>8.4%</td>
<td>-2.1</td>
<td>-0.3</td>
</tr>
<tr>
<td>Number of other ambulatory encounters (OA)</td>
<td>1.1</td>
<td>2.6%</td>
<td>1.3</td>
<td>3.5%</td>
<td>-0.4</td>
<td>-0.1</td>
</tr>
</tbody>
</table>

### Angioedema: Table 1. Unmatched Cohort

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>ACE Inhibitors</th>
<th>Beta Blockers</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patients</strong></td>
<td>2,211,215</td>
<td>1,673,682</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Events while on therapy</td>
<td>5,158</td>
<td>1,292</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Person-time at risk (days)</td>
<td>186.9</td>
<td>266.6</td>
<td>37.7</td>
<td>0.2</td>
</tr>
</tbody>
</table>

**Patient Characteristics**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ACE Inhibitors</th>
<th>Beta Blockers</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (F)</td>
<td>937,952</td>
<td>946,344</td>
<td>-11.4</td>
<td>-0.2</td>
</tr>
<tr>
<td>Mean age (std dev)</td>
<td>54.6</td>
<td>53.7</td>
<td>0.9</td>
<td>0.1</td>
</tr>
</tbody>
</table>

**Recorded History of**

<table>
<thead>
<tr>
<th>History</th>
<th>ACE Inhibitors</th>
<th>Beta Blockers</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allergic reactions</td>
<td>207,344</td>
<td>190,337</td>
<td>-7.0</td>
<td>-0.1</td>
</tr>
<tr>
<td>Diabetes</td>
<td>471,661</td>
<td>173,083</td>
<td>11.0</td>
<td>0.3</td>
</tr>
<tr>
<td>Heart failure</td>
<td>43,960</td>
<td>74,897</td>
<td>-2.6</td>
<td>-0.1</td>
</tr>
<tr>
<td>Ischemic heart disease</td>
<td>129,948</td>
<td>224,681</td>
<td>6.4</td>
<td>0.3</td>
</tr>
<tr>
<td>NSAID use</td>
<td>358,238</td>
<td>250,857</td>
<td>-0.6</td>
<td>0.0</td>
</tr>
</tbody>
</table>

**Health Service Utilization Intensity**

<table>
<thead>
<tr>
<th>Service</th>
<th>Mean (Std)</th>
<th>Mean (Std)</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of generics</td>
<td>5.4 (3.5)</td>
<td>3.5 (3.5)</td>
<td>0.9</td>
<td>0.1</td>
</tr>
<tr>
<td>Number of filled prescriptions</td>
<td>7.5 (9.6)</td>
<td>9.6 (9.6)</td>
<td>0.1</td>
<td>0.4</td>
</tr>
<tr>
<td>Number of inpatient hospital encounters (IP)</td>
<td>0.1 (0.4)</td>
<td>0.1 (0.4)</td>
<td>-0.1</td>
<td>-0.1</td>
</tr>
<tr>
<td>Number of non-acute institutional encounters (IS)</td>
<td>0.0 (0.6)</td>
<td>0.0 (0.6)</td>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td>Number of emergency room encounters (ED)</td>
<td>0.2 (0.7)</td>
<td>0.7 (0.7)</td>
<td>-0.1</td>
<td>-0.2</td>
</tr>
<tr>
<td>Number of ambulatory encounters (AV)</td>
<td>4.8 (6.3)</td>
<td>6.9 (8.4)</td>
<td>-2.1</td>
<td>-0.3</td>
</tr>
<tr>
<td>Number of other ambulatory encounters (OA)</td>
<td>1.1 (2.6)</td>
<td>1.5 (3.5)</td>
<td>-0.4</td>
<td>-0.1</td>
</tr>
</tbody>
</table>

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

Propensity Scores Before Match
## Angioedema: Table 2. Matched Cohort

**2.6 million new users**

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

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Primary Analysis</th>
<th>Covariate Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACE Inhibitors</td>
<td>Beta Blockers</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td><strong>%</strong></td>
<td><strong>N</strong></td>
</tr>
<tr>
<td>Patients</td>
<td>1,309,104</td>
<td>1,309,104</td>
</tr>
<tr>
<td>Events while on therapy</td>
<td>3,311</td>
<td>988</td>
</tr>
<tr>
<td>Person-time at risk (days)</td>
<td>183.8</td>
<td>151.8</td>
</tr>
</tbody>
</table>

### Patient Characteristics

<table>
<thead>
<tr>
<th></th>
<th>ACE Inhibitors</th>
<th>Beta Blockers</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (F)</td>
<td>723,955</td>
<td>689,617</td>
<td>55.3%</td>
<td>52.7%</td>
</tr>
<tr>
<td>Mean age (std dev)</td>
<td>54.1</td>
<td>54.4</td>
<td>13.1</td>
<td>14.0</td>
</tr>
</tbody>
</table>

### Recorded History of:

<table>
<thead>
<tr>
<th>Condition</th>
<th>ACE Inhibitors</th>
<th>Beta Blockers</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allergic reactions</td>
<td>157,920</td>
<td>154,933</td>
<td>10.5%</td>
<td>10.3%</td>
</tr>
<tr>
<td>Diabetes</td>
<td>150,056</td>
<td>150,551</td>
<td>11.5%</td>
<td>11.5%</td>
</tr>
<tr>
<td>Heart failure</td>
<td>38,302</td>
<td>38,966</td>
<td>2.7%</td>
<td>3.0%</td>
</tr>
<tr>
<td>Ischemic heart diseases</td>
<td>102,200</td>
<td>106,786</td>
<td>7.8%</td>
<td>8.2%</td>
</tr>
<tr>
<td>NSAID use</td>
<td>191,758</td>
<td>189,812</td>
<td>7.7%</td>
<td>14.3%</td>
</tr>
</tbody>
</table>

### Health Service Utilization Intensity:

<table>
<thead>
<tr>
<th>Service</th>
<th>Mean</th>
<th>Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of generics</td>
<td>3.7</td>
<td>3.7%</td>
</tr>
<tr>
<td>Number of filled prescriptions</td>
<td>8.1</td>
<td>10.2%</td>
</tr>
<tr>
<td>Number of inpatient hospital encounters</td>
<td>0.1</td>
<td>0.5%</td>
</tr>
<tr>
<td>Number of non-acute institutional encounters</td>
<td>0.1</td>
<td>0.7%</td>
</tr>
<tr>
<td>Number of emergency room encounters</td>
<td>0.3</td>
<td>0.8%</td>
</tr>
<tr>
<td>Number of ambulatory encounters</td>
<td>5.6</td>
<td>7.3%</td>
</tr>
<tr>
<td>Number of other ambulatory encounters</td>
<td>1.2</td>
<td>2.9%</td>
</tr>
</tbody>
</table>

### Diagnoses

- **Diabetes**: 10% vs 10%
- **Heart failure**: 3% vs 3%
- **Ischemic heart disease**: 8% vs 8%
Propensity Scores After Match

Histogram of PS distribution by DP (masked)
Histogram of Predefined PS among Predefined PS Matched Cohort, Matched Cal = .025  C-Stat for Predefined: 0.695

Angioedema: Table 3. Results

Table 3: Sequential Estimates for Angioedema Events by Analysis Type, and Drug Pair

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

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
---|---|---|---|---|---
ACE Inhibitors | 4.558 | 2.33 | 2.67 | 1.56 | 2.55 (2.40, 2.71) | <.0001
Beta Blockers | 1.890 | 0.77 | | | | |
ACE Inhibitors | 5.027 | 2.53 | 3.21 | 1.77 | 3.14 (2.86, 3.44) | <.0001
Beta Blockers | 1.815 | 0.75 | | | | |

Angioedema: Table 3. Results

ACEI vs β-blocker 1:1 matched analysis:

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

Table 3: Sequential Estimates for Angioedema Events by Analysis Type, and Drug Pair

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

Hazard Ratio (95% CI) and Wald P-Value:

<table>
<thead>
<tr>
<th>Incidence Rate per 1000 Person Years</th>
<th>Difference per 1000 Person Years</th>
<th>Difference in Risk per 1000 New Users</th>
<th>Hazard Ratio (95% CI)</th>
<th>Wald P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.558</td>
<td>2.33</td>
<td>2.55 (2.40, 2.71)</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>1.890</td>
<td>0.77</td>
<td>1.56</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>5.027</td>
<td>2.53</td>
<td>3.14 (2.86, 3.44)</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>1.815</td>
<td>0.75</td>
<td>1.77</td>
<td>&lt;.0001</td>
<td></td>
</tr>
</tbody>
</table>

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; >7 million 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; >7 million 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?
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?
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
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@pcpmednet.org

DRN Governance Document, v1.0
Distributed Research Network User’s Guide, v5.0 (PDF)
DRN Request Form (.docx)
Thank you!