The Sentinel System: the Case for Analysis Ready Data

Richard Platt MD, MS, and Jeff Brown, PhD.
Harvard Pilgrim Health Care Institute and Harvard Medical School
for the Sentinel Investigators
June 23, 2017
Learning Healthcare System

“The increased complexity of health care requires a sustainable system that gets the right care to the right people when they need it, and then captures the results for improvement. The nation needs a healthcare system that learns.”
- Data collected for one purpose aren’t reliably useful for other purposes
<table>
<thead>
<tr>
<th>Platelet count original result units³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blank</td>
</tr>
<tr>
<td>%</td>
</tr>
<tr>
<td>/100 W</td>
</tr>
<tr>
<td>/CMM</td>
</tr>
<tr>
<td>CMM</td>
</tr>
<tr>
<td>10 3 L</td>
</tr>
<tr>
<td>10X3UL</td>
</tr>
<tr>
<td>10^3/UL</td>
</tr>
<tr>
<td>10*3/uL</td>
</tr>
<tr>
<td>10?3/uL</td>
</tr>
<tr>
<td>10E3/uL</td>
</tr>
<tr>
<td>10e3/uL</td>
</tr>
<tr>
<td>10e9/L</td>
</tr>
<tr>
<td>E9/L</td>
</tr>
<tr>
<td>BIL/L</td>
</tr>
<tr>
<td>bil/L</td>
</tr>
<tr>
<td>CU MM</td>
</tr>
</tbody>
</table>
Data must be fit for intended purpose

- Accuracy
  - Identifying potential candidates for a clinical trial
  - Making a regulatory decision affecting a widely used drug
Data must be fit for intended purpose

- Accuracy
  - Identifying potential candidates for a clinical trial
  - Making a regulatory decision affecting a widely used drug
- Time required to be ready for analysis
Data must be fit for intended purpose

- **Accuracy**
  - Identifying potential candidates for a clinical trial
  - Making a regulatory decision affecting a widely used drug

- **Time required to be ready for analysis**

- **Need to use all data vs a subset**
Sentinel is a National Medical Product Monitoring System

LEARN MORE
Sentinel’s charge

Assess the use, safety, and effectiveness of regulated medical products by using electronic healthcare data plus other resources

Create data, informatics, and methodologic capabilities to support these activities
Sentinel partner organizations

Lead – HPHC Institute

Data and scientific partners

Scientific partners
Sentinel distributed database*

- Populations with well-defined person-time for which most medically-attended events are known
  - 223 million unique member IDs
  - 425 million person-years of observation time
  - 43 million people currently accruing new data
  - 5.9 billion dispensings
  - 7.2 billion unique encounters
  - 42 million people with >1 laboratory test result

* As of January 2017
Sentinel distributed analysis

1- User creates and submits query
2- Data Partners retrieve query
3- Data Partners review and run query against their local data
4- Data Partners review results
5- Data Partners return results via secure network
6- Results are aggregated and returned

https://www.sentinelinitiative.org/privacy-and-security
Three ways to address questions

Routine Analytic Framework (RAF)
- Off-the-shelf query “templates”
- Standard inputs, standard output
- Quick execution

RADaR: Rapid Analytic Development and Response:
- Hybrid approach: custom code leveraging RAF
- Standard inputs, custom output

Custom Programs
- Analysis as specified
- Custom inputs, custom output
- Longer execution
Selected protocol based assessments

- **CDER**
  - Dabigatran and several outcomes
  - Metabolic effects of 2\textsuperscript{nd} generation antipsychotics in youth
  - Diabetes drugs and acute myocardial infarction
  - IV Iron and anaphylaxis

- **CBER**
  - IV Immune Globulin and thromboembolic events
  - Gardasil and venous thromboembolism
  - Influenza vaccines and pregnancy outcomes
  - Gardasil 9 and Pregnancy Outcomes
  - Prevnar 13 and Kawasaki disease
  - Blood components and Transfusion-Related Lung Injury (TRALI)
Three ways to address questions

Rapid Analyses

Routine Analytic Framework (RAF)
- Off-the-shelf query "templates"
- Standard inputs, standard output
- Quick execution

RADaR: Rapid Analytic Development and Response:
RAF + custom code
- Hybrid approach: custom code leveraging RAF
- Standard inputs, custom output

Custom Programs
- Analysis as specified
- Custom inputs, custom output
- Longer execution
Sentinel’s tools

Cohort ID and Descriptive Analysis (CIDA) Tool
Options:
• Propensity Score Matching or Stratification
• Self-controlled Risk Interval Design
• Drug Use in Pregnancy
• Drug Utilization
• Concomitant Drug Utilization
• Pre/Post Index Tool
Routine Analytic Framework tools

- Validated, flexible, and reusable analytic programs
- Run efficiently against the Sentinel CDM and generate standardized output
- Optimized to meet FDA’s needs for responsiveness, data quality, reproducibility, and transparency
- Meets needs of Data Partners with diverse technical, data governance, security, and confidentiality requirements
Rapid analysis querying sequence

1. Simple counts (Summary tables)
2. Complex counts (Level 1)
3. Compare event rates (Level 2)
4. Follow-up (PEPR)

- Determine use and frequency
- Identify/describe population
- Comparative assessment
- New queries; Line Lists; Chart Review
Querying sequence

1. Simple counts
2. Complex counts
3. Compare event rates
4. Follow-up

- Determine use and frequency
- Identify/describe population
- Comparative assessment
- New queries; Line Lists; Chart Review
Simple counts (summary table queries)

- Counts of (new) users with exposure or condition
- Example: Dispensing of evolocumab (PCSK9 inhibitor) without prior dispensing during preceding 180 days, by age, sex, and year

<table>
<thead>
<tr>
<th>Age</th>
<th>2015</th>
<th>2016 (partial)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>&lt;44</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>45-64</td>
<td>85</td>
<td>61</td>
</tr>
<tr>
<td>65-74</td>
<td>42</td>
<td>35</td>
</tr>
<tr>
<td>75+</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>TOTAL</td>
<td>261</td>
<td></td>
</tr>
</tbody>
</table>

49 such queries / 291 scenarios in 2016

Querying sequence

Simple counts → Complex counts → Compare event rates → Follow-up

- Determine use and frequency
- Identify/describe population
- Comparative assessment
- New queries; Line Lists; Chart Review
Complex count queries (Level 1 / 1+)

- Counts and rates of events within user specified times, among populations identified using complex “and/or/not” relationships.
  - Example: Rates of first diagnosis of heart failure or cardiomyopathy among new users of different drugs used to treat ADHD, by age and duration of exposure

- 53 queries, 800+ scenarios in 2016
Figure 1. Rate of heart failure events (per 10,000 person years) by age group, medication, and duration of use

Heart failure events per 10,000 person-years

Duration of use in days and age group

<22 Years 22-44 Years 45-64 Years 65+ Years
0-90 181-270 366-730 731-1095
181-270 271-365 366-730 731-1095
366-730 731-1095

- Amphetamine
- Methylphenidate
- Atomoxetine
Complex count queries (Level 1 / 1+)

• Counts and rates of events within user specified times, among populations identified using complex “and/or/not” relationships.
  – Example: Rates of first diagnosis of heart failure or cardiomyopathy among new users of different drugs used to treat ADHD, by age and duration of exposure

• 53 queries, 800+ scenarios in 2016

• New uses
  – Medications errors (name confusion, dosing errors)
  – Geographic location stratification
Sentinel’s tools

Summary Table Tool

Cohort ID and Descriptive Analysis (CIDA) Tool

Options:
- Propensity Score Matching or Stratification
- Self-controlled Risk Interval Design
- Drug Use in Pregnancy
- Drug Utilization
- Concomitant Drug Utilization
- Pre/Post Index Tool
Use of antiemetic drugs among live birth pregnancies in the Sentinel Distributed Database, 2001-2014\textsuperscript{a,b}

- Any antiemetic use
- Ondansetron - oral
- Ondansetron - injectable
- Doxylamine/Pyridoxine
- Metoclopramide
- Promethazine

\textsuperscript{a} Dashed lines for oral and injection ondansetron form represent a portion of all total ondansetron use as shown by the solid purple line. Summation of oral and injection utilization sums to greater than total ondansetron use since some women received both products.

\textsuperscript{b} Not all Mini-Sentinel data partners contributed data for the entire study period.

Taylor. Pharmacoepidemiology and Drug Safety 2017;26:592
Recent urgent request

- Issue related to concomitant drug use
- Two similar drugs
  - Drug A has known interaction with Drug Class X
  - Drug B does not have known interaction with Drug Class X
- Goal: Estimate the proportion of concomitant use of Drug A and Drug Class X compared to proportion of concomitant use of Drug B and Drug Class X
Query Timeline

May 31  June 1  June 2  June 3  June 4  June 5
Query Timeline: May 31

12:33pm: Operations Center receives URGENT query request

By 2pm: Teleconference with FDA to discuss the request

4:20pm: First draft of query specifications ready for internal review

By end of day: Specifications ready for FDA review
30

10am: Operations center reviews draft specifications with FDA

11am: Content expert at Data Partner responds to email offering additional support and consultation

1:38pm: First draft request package assembled (specifications not finalized yet)

3:16pm: FDA request modification to specifications

4pm: Consult with content expert (at Data Partner site)

5:10pm: Final specifications sent to FDA for approval

By end of day: Revised query package ready for quality assurance and scientific review checks
By 9:45am: Query package passed technical and scientific review and quality assurance measures

10:21am: Query package distributed to 16 Sentinel Data Partners

11:01am: First set of results arrive from Data Partners

By end of day: 12 of 16 Sentinel Data Partners have returned results

8:47am: FDA approves query specifications
Query Timeline: June 3 & 4

June 3 & 4: It’s the weekend!
Results are aggregated and initial report is drafted and ready for review at Operations Center.

2:58pm: Last Sentinel Data Partner has uploaded query results.

4:23pm: Final report passed internal technical and scientific review.

4:36pm: Report sent to FDA.

By 10:30am: 15 of 16 Data Partners have responded to data request.
Query Timeline = 3.5 business days
Querying sequence

1. Simple counts
2. Complex counts
3. Compare event rates
4. Follow-up

- Determine use and frequency
- Identify/describe population
- Comparative assessment
- New queries; Line Lists; Chart Review
Comparison of rates (Level 2 / 2+)

- Adjusted relative rates or hazard ratios comparing outcomes among two cohorts identified by complex count program
  - or
- Adjusted self-controlled risk interval analysis
  - Example: Risk of seizures associated with new use of ranolazine
- 11 queries / 100+ scenarios in 2016
Sentinel’s tools

Summary Table Tool

Cohort ID and Descriptive Analysis (CIDA) Tool

Options:
- Propensity Score Matching or Stratification
- Self-controlled Risk Interval Design
- Drug Use in Pregnancy
- Drug Utilization
- Concomitant Drug Utilization
- Pre/Post Index Tool
## Angioedema: Table 1. Unmatched cohort

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

<table>
<thead>
<tr>
<th>Character</th>
<th>Inhibitors</th>
<th>Beta Blockers</th>
<th>Primary Analysis</th>
<th>Covariate Balance</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>Character</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (F)</td>
<td>997,962</td>
<td>45.10%</td>
<td>946,344</td>
<td>56.50%</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>

### Recorded History of:

<table>
<thead>
<tr>
<th>Event</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anaphylactic reactions</td>
<td>207,344</td>
<td>9.4%</td>
<td>190,587</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>173,083</td>
<td>10.3%</td>
<td>11.0</td>
<td>0.3</td>
</tr>
<tr>
<td>Heart failure</td>
<td>41,060</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 disease</td>
<td>109,948</td>
<td>5.0%</td>
<td>224,081</td>
<td>15.4%</td>
<td>-8.4</td>
<td>-0.3</td>
</tr>
<tr>
<td>NSAID use</td>
<td>318,298</td>
<td>5.0%</td>
<td>250,667</td>
<td>15.0%</td>
<td>-6.6</td>
<td>0.0</td>
</tr>
</tbody>
</table>

### Health Service Utilization Intensity:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of generics</td>
<td>3.4</td>
<td>3.5</td>
</tr>
<tr>
<td>Number of filled prescriptions</td>
<td>7.5</td>
<td>9.6</td>
</tr>
<tr>
<td>Number of inpatient hospital encounters (IP)</td>
<td>0.1</td>
<td>0.4</td>
</tr>
<tr>
<td>Number of non-acute institutional encounters (IS)</td>
<td>0.0</td>
<td>0.6</td>
</tr>
<tr>
<td>Number of emergency room encounters (ED)</td>
<td>0.2</td>
<td>0.7</td>
</tr>
<tr>
<td>Number of ambulatory encounters (AV)</td>
<td>4.8</td>
<td>6.3</td>
</tr>
<tr>
<td>Number of other ambulatory encounters (OA)</td>
<td>1.1</td>
<td>2.6</td>
</tr>
</tbody>
</table>

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

3.9 million new users
### Angioedema: Table 2. Matched cohort

#### 2.6 million new users

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>E Inhibitors</th>
<th>Beta Blockers</th>
<th>Absolute Difference</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>winters</td>
<td>1,309,104</td>
<td>1,309,104</td>
<td>0.0</td>
<td>-0.4</td>
</tr>
<tr>
<td>Events while on therapy</td>
<td>3,311</td>
<td>988</td>
<td>0.2</td>
<td>0.0</td>
</tr>
<tr>
<td>Person-time at risk (days)</td>
<td>183.8</td>
<td>151.8</td>
<td>31.9</td>
<td>0.1</td>
</tr>
<tr>
<td>Gender (F)</td>
<td>723,955</td>
<td>689,617</td>
<td>2.6</td>
<td>0.1</td>
</tr>
<tr>
<td>Mean age (std dev)</td>
<td>54.1</td>
<td>54.4</td>
<td>-0.3</td>
<td>0.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Recorded History of:</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Allergic reactions</td>
<td>137,920</td>
<td>134,933</td>
<td>0.2</td>
<td>0.0</td>
</tr>
<tr>
<td>Diabetes</td>
<td>150,036</td>
<td>150,551</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Heart failure</td>
<td>35,302</td>
<td>38,966</td>
<td>-0.3</td>
<td>0.0</td>
</tr>
<tr>
<td>Ischemic heart disease</td>
<td>102,200</td>
<td>106,786</td>
<td>-0.4</td>
<td>0.0</td>
</tr>
<tr>
<td>NSAID use</td>
<td>191,798</td>
<td>189,612</td>
<td>0.2</td>
<td>0.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Health Service Utilization Intensity:</th>
<th>Mean</th>
<th>Std Dev</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</td>
<td>0.1</td>
<td>0.5%</td>
</tr>
<tr>
<td>encounters (IP)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of non-acute institutional</td>
<td>0.1</td>
<td>0.7%</td>
</tr>
<tr>
<td>encounters (IS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of emergency room encounters (ED)</td>
<td>0.3</td>
<td>0.8%</td>
</tr>
<tr>
<td>Number of ambulatory encounters (AV)</td>
<td>5.6</td>
<td>7.3%</td>
</tr>
<tr>
<td>Number of other ambulatory encounters (OA)</td>
<td>1.2</td>
<td>2.9%</td>
</tr>
</tbody>
</table>

- **Diabetes**: 10% vs 10%
- **Heart failure**: 3% vs 3%
- **Ischemic heart disease**: 8% vs 8%
Angioedema: Table 3. Results

ACEI vs β-blocker 1:1 matched analysis:
• HR = 3.1 (95% CI, 2.9-3.4)

Toh et al findings:
• HR = 3.0 (95% CI, 2.8-3.3)
Querying sequence

1. Simple counts
2. Complex counts
3. Compare event rates
4. Follow-up

Detailed steps:
- Determine use and frequency
- Identify/describe population
- Comparative assessment
- New queries; Line Lists; Chart Review
Patient Episode Profile Retrieval (PEPR)

Day 0, office visit
Routine health check
Immunization

Day 4, office visit
Gastroenteritis

Day 7, hospitalized
Vomiting / cough
Dehydration
Gastroenteritis

---

### Episode Detail

<table>
<thead>
<tr>
<th>Days from expos</th>
<th>Enc type</th>
<th>LOS</th>
<th>Clinical code</th>
<th>Code description</th>
<th>Incidence^</th>
<th>P Dx#</th>
<th>Node (Y/N)</th>
<th>Main expos (Y/N)</th>
<th>Any vacc (Y/N)</th>
<th>Rx days supp</th>
<th>Rx amt</th>
<th>Cov start~</th>
<th>Cov end~</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>AV</td>
<td>DX</td>
<td>09 V0382</td>
<td>Need Proph Vacc Agnst Strep Pne</td>
<td>F</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>0</td>
<td>AV</td>
<td>DX</td>
<td>09 V068</td>
<td>Need Proph Vacc Against Oth Comb Dz</td>
<td>F</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>0</td>
<td>AV</td>
<td>DX</td>
<td>09 V202</td>
<td>Routine Infant/Child Health Check</td>
<td>F</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>0</td>
<td>AV</td>
<td>PX</td>
<td>C4 90471</td>
<td>Immunization Admin</td>
<td>F</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>0</td>
<td>AV</td>
<td>PX</td>
<td>C4 90472</td>
<td>Immunization Admin Each Add</td>
<td>F</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>0</td>
<td>AV</td>
<td>PX</td>
<td>C4 90669</td>
<td>PCV7 Vaccine Im</td>
<td>F</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>0</td>
<td>AV</td>
<td>PX</td>
<td>C4 90710</td>
<td>MMRV Vaccine Sc</td>
<td>F</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>0</td>
<td>AV</td>
<td>PX</td>
<td>C4 99392</td>
<td>Prev Visit Est Age 1-4</td>
<td>F</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>4</td>
<td>AV</td>
<td>DX</td>
<td>09 0090</td>
<td>Inf Colitis Enterit &amp; Gastroenterit</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>4</td>
<td>AV</td>
<td>PX</td>
<td>C4 99213</td>
<td>Office/Outpatient Visit Est</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>7</td>
<td>IP</td>
<td>DX</td>
<td>09 27651</td>
<td>Dehydration</td>
<td>I</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>7</td>
<td>IP</td>
<td>DX</td>
<td>09 53550</td>
<td>Uns Gastrit &amp; Gastroduodit No Hemorr</td>
<td>I</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>7</td>
<td>IP</td>
<td>DX</td>
<td>09 7862</td>
<td>Cough</td>
<td>I</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>7</td>
<td>IP</td>
<td>DX</td>
<td>09 78703</td>
<td>Vomiting Alone</td>
<td>I</td>
<td>S</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>7</td>
<td>IP</td>
<td>PX</td>
<td>C4 71020</td>
<td>Chest X-Ray Ziiew Frontal &amp; Lat</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
<tr>
<td>7</td>
<td>IP</td>
<td>PX</td>
<td>C4 74000</td>
<td>X-Ray Exam Of Abdomen</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-386</td>
<td>1260</td>
</tr>
</tbody>
</table>

^ Incidence: F = first observed; I = incident; blank = prevalent
# Primary Dx: P = primary; S = secondary; X = N/A
~ Med enroll segment containing the admission date of the encounter or the drug enroll segment containing the dispensing date
In theory there is no difference between theory and practice. In practice there is.

Yogi Berra

www.brainyquote.com/quotes/quotes/y/yogiberra141506.html#gsD0IBx3dytirLPX.99
How does it work?

Routine Tools combined with Robust Data Quality Assurance Practices
Every Data Partner transforms their data into the Sentinel Common Data Model.
The quality assurance process

Send a standard QA checking program to check DP’s ETL in waiting

**Compliance Checks**
- **Level 1**: Completeness, validity, accuracy
- **Level 2**: Cross-variable and cross-table integrity

**Judgment Call Checks**
- **Level 3**: Trends: consistency
- **Level 4**: Logical: plausibility, convergence

Data Partner
QA example: Admission and discharge date

Completeness
- Admission date (ADate) variable has missing values

Validity
- ADate variable is not SAS date value of numeric data type
- ADate variable is not of length 4

Accuracy
- ADate is before DDate (for IP and IS only)
- ADate and DDate variables have values after DP_MinDate

Integrity
- Discharge date (DDate) variable is missing for EncType value "IP"
- DDate variable is populated for EncType values other than "IP" or "IS"

*IP = Inpatient Setting, IS= Institutional Setting like a Skilled Nursing Facility
The quality assurance process

Send a standard QA checking program to check DP’s ETL in waiting

Compliance Checks
Level 1: Completeness, validity, accuracy
Level 2: Cross-variable and cross-table integrity

Judgment Call Checks
Level 3: Trends: consistency
Level 4: Logical: plausibility, convergence

Data Partner
The database is dynamic – updates overwrite the preceding data!

**Data Delivery 1**

**Data Delivery 2**

---

**Unique Data Partner Source Database Structure**

**Data Partner’s Database Transformed into SCDM Format**

**Timeframe of Data Available in Database**

QA example: Admission / discharge dates

Check distributions and patterns for notable changes

Consistency

• Problem with distribution of ADate (i.e. total number of records per year) within the ETL
• Problem with distribution of ADate (i.e. total number of records per year-month) within the ETL
• Significant change in number of records per ADate (year) across ETLs
• Significant change in number of records per ADate (year-month) across ETLs
• Problem with distribution of ADate (overall) within the ETL
• Problem with distribution of ADate (overall) across ETLs
• Problem with distribution of DDate variable by EncType per year-month
• Problem with distribution of length of stay by EncType per year
Sentinel war stories: Consistency checks

Is source of inconsistency clear error or Data Partner changes / improvements?

Incorrect Data Load

Reclassification of Encounter Type
Sentinel QA statistics

- Annually, the QA team conducts reviews for approximately 50 data deliveries from 17 Data Partners

- Since 1/1/2016, the QA package has had to be re-run in 16 instances to fix an issue

- In the latest data deliveries from the 5 largest DPs, 25 checks were reported in QA that required DP follow-up
  - 22 of the 25 were Level 3 checks
Data Review Tool – Account of Issues

1) Select MSCDM Table: Diagnosis

2) Choose Error Source Dataset To Evaluate: dia_13_pdx_et

3) Select Data Check To Evaluate: DIA2.9.2

4) Data Check Description:
   PDX variable is populated for EncType values other than “IP” or “IS”

5) Data Check Evaluation Guideline:
   Should be populated for IP and IS and missing for ED, AV, and OA EncTypes.
   Review inappropriately populated values of PDX. If values are populated for EncTypes other than “IP” or “IS”, review previous ETL reports to see if the issue has been raised and addressed before. If it has, and MSOC has agreed to allow PDX to be populated for EncTypes other than “IP” or “IS”, document this issue on the current ETL report and note that we will continue to monitor. If the issue has not been addressed before, ask the DP to investigate and explain.

6) Describe Issue Here If Check Fails

7) Select Pass, Fail, or Ignore For Selected Check:
   - PASS
   - FAIL
   - IGNORE

Problems viewing the dataset? Click on this button to go to the dataset’s worksheet
   Go To Dataset

8) Click this button to check if you have completed the review
   - CORE
   - LABS

53
Sentinel operations center Quality Assurance team

- Refreshes per Year: ~50
- 1 Manager, 2 Programmers, 3 Analysts
- Tasks:
  - Oversight
  - Maintenance and troubleshooting
  - Updating and distribution of quality assurance programs
  - Aggregation and reporting
Lab Data requires more extensive QA support

Electronic clinical laboratory test results data tables: lessons from Mini-Sentinel

Marsha A. Raebel\textsuperscript{1,2*}, Kevin Haynes\textsuperscript{3}, Tiffany S. Woodworth\textsuperscript{4}, Gwyn Saylor\textsuperscript{5}, Elizabeth Cavagnaro\textsuperscript{4}, Kara O. Coughlin\textsuperscript{4}, Lesley H. Curtis\textsuperscript{6}, Mark G. Weiner\textsuperscript{7,8}, Patrick Archdeacon\textsuperscript{8} and Jeffrey S. Brown\textsuperscript{4}

\textsuperscript{1}Institute for Health Research, Kaiser Permanente Colorado, Denver, CO, USA
\textsuperscript{2}University of Colorado Skaggs, School of Pharmacy and Pharmaceutical Sciences Aurora, CO, USA
\textsuperscript{3}Center for Clinical Epidemiology and Biostatistics and Center for Pharmacoepidemiology Research and Training, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
\textsuperscript{4}Department of Population Medicine, Harvard Pilgrim Health Care Institute and Harvard Medical School, Boston, MA, USA
\textsuperscript{5}Kaiser Permanente Northwest Center for Health Research, Portland, OR, USA
\textsuperscript{6}Duke Clinical Research Institute and Department of Medicine, Duke University School of Medicine, Durham, NC, USA
\textsuperscript{7}Department of Clinical Sciences, Temple University School of Medicine, Philadelphia, PA, USA
\textsuperscript{8}Center for Drug Evaluation and Research, Food and Drug Administration Silver Spring, MD, USA
## Variations in result units in source data

### Platelet count original result units

<table>
<thead>
<tr>
<th>Unit</th>
<th>Original Unit</th>
<th>Conversion Unit</th>
<th>Conversion Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blank</td>
<td>FL</td>
<td>TH/UL</td>
<td>X10(3)</td>
</tr>
<tr>
<td>%</td>
<td>K/CMM</td>
<td>THOU/CMM</td>
<td>1000/UL</td>
</tr>
<tr>
<td>/100 W</td>
<td>k/cmm</td>
<td>thou/cmm</td>
<td>X10(3)/MCL</td>
</tr>
<tr>
<td>/CMM</td>
<td>K/CU MM</td>
<td>thou/mm3</td>
<td>X10(3)/UL</td>
</tr>
<tr>
<td>CMM</td>
<td>K/CUMM</td>
<td>THOU/UL</td>
<td>X10(6)/MCL</td>
</tr>
<tr>
<td>10 3L</td>
<td>K/MCL</td>
<td>THOUS/CU.MM</td>
<td>X10*9/L</td>
</tr>
<tr>
<td>10X3UL</td>
<td>K/mcL</td>
<td>THOUS/MCL</td>
<td>X10E3/UL</td>
</tr>
<tr>
<td>10^3/UL</td>
<td>K/UL</td>
<td>THOU/mcL</td>
<td>X1000</td>
</tr>
<tr>
<td>10*3/uL</td>
<td>k/uL</td>
<td>THOUS/UL</td>
<td>X10X3</td>
</tr>
<tr>
<td>10^3/uL</td>
<td>KU/L</td>
<td>Thou/uL</td>
<td>X10^3/UL</td>
</tr>
<tr>
<td>10E3/uL</td>
<td>K/MM3</td>
<td>THOUSA</td>
<td>x10</td>
</tr>
<tr>
<td>10e3/uL</td>
<td>K/mm3</td>
<td>THOUSAND</td>
<td>X10^3/ul</td>
</tr>
<tr>
<td>10e9/L</td>
<td>LB</td>
<td>THOUSAND/UL</td>
<td>X10E3</td>
</tr>
<tr>
<td>E9/L</td>
<td>PLATELET CO</td>
<td>U</td>
<td>X10E3</td>
</tr>
<tr>
<td>BIL/L</td>
<td>T/CMM</td>
<td>X 10-3/UL</td>
<td>K/A?L</td>
</tr>
<tr>
<td>bil/L</td>
<td>TH/MM3</td>
<td>X 10(3)/UL</td>
<td>K/B5L</td>
</tr>
<tr>
<td>CU MM</td>
<td>th/mm3</td>
<td>X10 3</td>
<td></td>
</tr>
</tbody>
</table>
National Medical Evidence Generation Collaborative (EvGen Collaborative)

Resources for You

- Office of Medical Products and Tobacco

www.fda.gov/ScienceResearch/SpecialTopics/EvGenSystem
Coordinating Center(s)

Quality of Care
Sponsor(s)

Public Health Surveillance

Medical Product Safety Surveillance
FDA

Comparative Effectiveness Research

Payers
- Public
- Private

Providers
- Hospitals
- Physicians
- Integrated Systems

Registries
- Disease-specific
- Product-specific

Common Data Model
Data Standards

Queries

Results

CDC

Clinical Research

DISTRIBUTED NETWORK GOVERNANCE

Queries

Results

Coordinating Center(s)

Payers

Registries

Providers

Common Data Model
Data Standards

Comparative Effectiveness Research

Impact-AFib

Clinical Trials Transformation Initiative

NIH Collaboratory Distributed Research Network

NIH Collaboratory

NIH

Sentinel

Sentinel Center

Innovation in Medical Evidence Development and Surveillance

Biologics & Biosimilars Collective Intelligence Consortium

Randomized Clinical Trials

Clinical Research

Results

Results

Results

Results
Thank you!
### Summary of query specifications: Overall

1. **Select the Query Type (Level):**
   - Level 2: Cohort Selection and Analytic Adjustment

2. **Select the Analysis Tool:**
   - Propensity Score Matching Tool

3. **Describe Study Objectives:**
   - To assess the ability of Mini-Sentinel comparative assessment modular programs to reproduce the known association between ACEIs and angioedema

4. **Define Study Period:**
   - 01/01/2008 - 09/30/2013

5. **List the age group(s) of interest:**
   - 18 +

6. **Specify enrollment requirements:**
   - **Coverage type:** Medical and drug coverage
   - **Maximum enrollment gap (days):** 45
   - **Continuous enrollment before exposure (days):** 183

---

## Summary of query specifications: Exposures

<table>
<thead>
<tr>
<th>Exposures of Interest</th>
<th>Comparator of Interest (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACE inhibitors (benazepril, captopril, enalapril, fosinopril, lisinopril, moexipril, quinapril, perindopril, ramipril, or tranolapril)</td>
<td>Beta-blockers (acebutolol, atenolol, bisoprolol, carvedilol, labetalol, metoprolol, nebivolol, pindolol, propranolol, or timolol)</td>
</tr>
</tbody>
</table>

1. **Define exposures (generic/brand names):**  
   - ACE inhibitors (benazepril, captopril, enalapril, fosinopril, lisinopril, moexipril, quinapril, perindopril, ramipril, or tranolapril)  
   - Beta-blockers (acebutolol, atenolol, bisoprolol, carvedilol, labetalol, metoprolol, nebivolol, pindolol, propranolol, or timolol)

2. **Define exposure incidence:**  
   - Washout period (days): 183  
   - Beta-blockers, aliskiren, ARBs (candesartan, eprosartan, irbesartan, losartan, olmesartan, telmisartan, or valsartan)  
   - ACE inhibitors, aliskiren, ARBs (candesartan, eprosartan, irbesartan, losartan, olmesartan, telmisartan, or valsartan)

3. **Specify exposed time assessment (AT or ITT):**  
   - As Treated (AT)

4. **Specify follow-up duration (for ITT assessments; in days):**  
   - Leave blank for AT assessments
Summary of query specifications:

Additional information

- **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
Propensity scores before match

Histograms of PS distribution by DP (masked)
Histogram of Predefined PS, Unmatched Cohort  C-Stat for Predefined:  0.695

Density

PS

Normal Plot of ace_i  Normal Plot of beta_b  Histogram of ace_i  Histogram of beta_b

0.0  0.2  0.4  0.6  0.8  1.0

0  2  4  6  8  10

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
Sentinel’s tools

Cohort ID and Descriptive Analysis (CIDA) Tool

Options:
- Propensity Score Matching or Stratification
- Self-controlled Risk Interval Design
- Drug Use in Pregnancy
- Drug Utilization
- Concomitant Drug Utilization
- Pre/Post Index Tool
New user cohort design

- Look back XX days
- Inclusion/exclusion condition and/or treatment
- Start of **new** treatment episode
- Optional: blackout days
- Optional: extension days
- Outcome(s)

Start Date

End Date

Index Date

Time
Blood transfusion during pregnancy

- Need for rapid assessment of frequency of transfusion during pregnancy
- Sentinel Distributed Dataset identified 1,946,032 deliveries with coverage during entire pregnancy from 2008-2015 (~8% of U.S. deliveries)
- 21,048 (1.1%) pregnancies had blood transfusion
- Report with integrated data from 15 data partners returned to FDA within 3 working days of final specification

www.sentinelinitiative.org/vaccines-blood-biologics/assessments/blood-transfusions