Clinical Integrated Data Repositories and Governing Data Sharing

Russ Waitman, PhD, University of Kansas Medical Center
Director of Medical Informatics, Associate Vice Chancellor Enterprise Analytics
Professor, Internal Medicine
Greater Plains Collaborative Principal Investigator
May 24, 2017

This work is supported in part by NIH grant UL1TR000001 and the PCORI PCORnet Greater Plains Collaborative CDRN
Acknowledgements and Conflicts

This presentation brings together slides from many contributors including but not limited to PCORnet’s master slides introduction, the ADAPTABLE trial team, the Data Standards Security and Network Infrastructure Task Force, Dr. Robert Califf’s December presentation to CTSA PIs, and the Greater Plains Collaborative team. Contact me for further information: rwaitman@kumc.edu

I am a board member of the i2b2 tranSMART Foundation
Overview of This Talk

Part 1, Clinical Integrated Data Repositories common to United States Academic Medical Centers with National Institutes of Health funded Clinical and Translational Science Institutes.

Part 2, PCORnet and the Greater Plains Collaborative leveraging NIH CTSA investments

- Data Integration to serve Comparative Effectiveness Research
- How does a AMC keep order in it’s data kitchen?
- ADAPTABLE aspirin pragmatic trial: first prospective trial

Part 3, What kind of national network will emerge and how do you fit in?
Backstory: General Medical Informatics Hypotheses

Hypothesis #1: Admin + Clinical -> Better Knowledge?

Pre-encounter factors: typically not in electronic form accessible to provider

Environment
- "toxic sludge worker"

Social Behavioral
- "exercises by jogging 2/week"

Genetic
- "poor warfarin metabolizer"

Diet, drugs
- "carrots, alcohol, BetaBlocker"

Illness/complaint
- Assaults health
- Bad enough to make appt

Healthcare Encounter
- "Black box"
- Restores health
- Body's control system + time

Hypothesis #2: Computer + Clinical Process-> Better Health?

Vitals, assessments, diagnosis, pre-encounter factors (history, diagnosis, death)

Outcomes
- "Health" or physiologic reserve

"Health" or physiologic reserve

Observations (not recorded in EMR)

Observations about EMR use

Clinical Decision Support & Information Synthesis Tools

Decisions to do something $d_1$

Actions (not recorded in EMR)

Observations about CDS effectiveness $\theta_2$

Interventions $i_1$

Collect more data: Order Tests

Outcomes (directly measured by health care delivery system)

Outcomes (not typically measured)

Ex: Laboratory, Radiology, Pathology/ exploratory procedures

Medications (orders, prescriptions)

Procedures (surgery, PT/OT, amelioration)

Unregulated changes (diet)

Administrative billing activities collect "signals" as charges in parallel with clinical activities

Charges gleaned from EMR and paper $c_1$

Billing system (technical charges)

Billing system (professional charges)
Part 1, 2010 arrived in Kansas for CTSA Program
Developing Capacity for Biomedical and
Informatics Research: Keep it simple.. Data

Biomedical Informatics Can Help Your Research

- We have tools and expertise to manage data and convert it into information

- REDCap – enter and manage data

- HERON (i2b2) – fish for data from the hospital/clinic

- Biweekly Frontiers Clinical Informatics Clinics
  - Tuesday 12-1:30 pm in 3001D Student Center.
  - Next session May 2nd …. Free Pizza!
You’re that fisherman: wanting to land data to answer your research hypothesis

Bennett Spring Trout Park, Lebanon Missouri
http://mdc.mo.gov/regions/southwest/bennett-spring
The Fish: Diagnoses, Demographics, Observations, Treatments
I want to go fishing, not fill a fish tank (REDCap)
Use HERON: a managed fishery
Get a License: Develop business agreements, policies, data use agreements and oversight.

Get a Fishing Rod and Bass Boat: Implement open source NIH funded (i.e. i2b2 https://www.i2b2.org/) initiatives for accessing data.

Know what your catching: Transform data into information using the NLM UMLS Metathesaurus as our vocabulary source.
- Secondary goal; mostly irrelevant at one site

Stock Different Tasty Fish: link clinical data sources to enhance their research utility.
• Fill out System Access Agreements to sponsor students/staff
• Fill out Data Use Agreement to request data export
• No Limit!!! IRB Protocol Not Required to view or pull de-identified data

• Must be on campus or use VPN or https://access.kumed.com
• Check http://frontiersresearch.org/frontiers/HERON-Introduction for more information, status, and training videos

Single sign-on using your email username
Real-time check for current human subjects training
The i2b2 “Fishing Rod”: build Diabetes cohort

Types of “fish” in folders

Drag concepts from upper left into panels on the right
i2b2: **AND** in Frontiers Research Registry

Dragging over the second condition
i2b2: **AND** a high Hemoglobin A1C

When you add a numeric concept, i2b2 asks if you want to set a constraint.
i2b2 Result: **497** patients in Cohort

**Run the Query**  
Query took 4 seconds  
497 patient in cohort
i2b2: Explore Cohort, Visualize
The dream: landing the big one

http://www.oregon.com/columbia_gorge_attractions/bonneville_hatchery
Without getting bit

!! CAUTION !!

TROUT BITE

DO NOT PUT YOUR HANDS OR FINGERS IN THE WATER
HERON Integrated Data Repository Architecture

Participants Clinical Systems (EPIC, IDX, VELOS)

Information in files from Source Systems

Extract, Transform & Load Processes or HL7 Listeners

Identified staging database (Night HERON)

De-identified data server

De-identified staging database (Blue HERON)

I2b2 clinical business intelligence application (JBoss, VMWare virtualized host managed by KUMC Information Resources)

Clinical/Translational Researcher

QI i2b2

QI analyst/provider
Dedicated Coordinator/Honest Broker: Tamara McMahon. Informatics Clinics held biweekly and one-to-one trainings and consultations offered.

Integrating HERON’s use into other research workflows:

- **Finding patients for prospective trials:** combining the Frontiers Participant Registry with the EMR data to find willing participants that meet study criteria.
- **Searching for samples:** Biospecimen Repository combined with EMR to find tissues that meet research criteria.

Auditing small queries.
Part 2, PCORNet and Greater Plains Collaborative

Long term CTSA Aims: T2-T4 translations and support population level research

HERON seems to work… what’s next?

Portions of the PCORnet standard introduction

Greater Plains Collaborative introduction and approaches
Our national clinical research system is well-intentioned but flawed

- High percentage of decisions not supported by evidence*
- Health outcomes and disparities are not improving
- Current system is great except:
  - Too slow
  - Too expensive
  - Unreliable
  - Doesn’t answer questions that matter most to patients
  - Unattractive to clinicians & administrators

We are not generating the evidence we need to support the healthcare decisions that patients and their doctors have to make every day.

Both researchers and funders now recognize the value in integrating clinical research networks

- Linking existing networks means clinical research can be conducted more effectively
- Ensures that patients, providers, and scientists form true “communities of research”
- Creates “interoperability” – networks can share sites and data
PCORnet’s goal

PCORnet seeks to improve the nation’s capacity to conduct clinical research by creating a large, highly representative, national patient-centered network that supports more efficient clinical trials and observational studies.
PCORnet embodies a “community of research” by uniting systems, patients & clinicians

PCORnet: A national infrastructure for patient-centered clinical research
The Greater Plains Collaborative Partners

KS, the University of Kansas Medical Center (KUMC)

MO, Children’s Mercy Hospital, University of Missouri

IA, University of Iowa Healthcare

IN, Indiana University

WI, the University of Wisconsin-Madison, the Medical College of Wisconsin, and Marshfield Clinic

MN, the University of Minnesota Medical Center

NE, the University of Nebraska Medical Center

TX, the University of Texas Health Sciences Center at San Antonio and the University of Texas Southwestern Medical Center.
GPC Size, Goals and Structure in Phase 1

- 11.8 Million Covered Lives
- 13 hospitals, 430 clinics, 1800 primary care providers, 7600 specialists
- Establish Governance
- Measure EHR Meaningful Use standardization and align for 3 use cases:
  - Breast Cancer
  - ALS (Lou Gerhig’s Disease)
  - Obesity (Pediatric Inpatient Focus)
- Develop Patient Reported Outcome Measure Methods
- Develop Comparative Effectiveness Research Trial infrastructure embedded in EHRs
- Enhance Patient Recruitment
- i2b2 and REDCap technologies
The GPC held a “Boot Camp” the day before the sessions with GPC investigators to prepare the patients, patient advocates, and other community representatives in attendance to participate more fully in Cohort working group discussions.

Kim Kimminau and Cheryl Jernigan leading engagement

http://www.gpcnetwork.org/?q=LEK2014
http://www.gpcnetwork.org/?q=LEK2015
http://gpcnetwork.org/?q=2016LEC
Gross Observation:

**Trialists/Clinicians Want “Their” Rich Data:**

- Cancer Registry
- Morphology,
- Death Index,
- Patient Portal Usage
More Rich Data: Cardiovascular Registry (NCDR) – Arterial Stenosis
Rich Data:

University HealthSystem Consortium Visit Details,

EMR Visit Vitals
Rich Data: REDCap Patient Reported Outcome Measures for Alzheimer's Disease Center
Observationalists (Epidemologists) Want “The” Data

Create a research-ready dataset of at least 1 million patients that is:

- **Comprehensive**, using data from EHRs to describe patients’ care experience over time and in different care settings
- **Utilize** multiple rich data sources to support research, such as electronic health records, insurance claims data, and data reported directly by patients

Execution and Governance:
- Developing relationship with external data partners (CMS, State, private insurers)

**Figure 3.1.** Comprehensive and complete data example from KUMC: heat map of percentage of proposed data elements from the HER and billing sources recorded in six month intervals surrounding the data of breast cancer diagnosis specified by the hospital tumor registry.
Everything’s Amazing and Nobody’s Happy\textsuperscript{1} and... it’s the Celebrity Chef’s fault....

- **Informatician:** you wanted anchovies for your pizza; I got your anchovies, pal. You no longer have to get residents to catch them by hand at night with a flashlight.

- **Researcher:** But I want to use Guy Fieri’s\textsuperscript{2} recipe for my study design... he’s so charismatic and spellbinds study sections and journal editors! We can’t be in the XYZ pizza making initiative if we don’t.

- **Guy Fieri:** My pizza recipe is the best but only works if you construct an oven to my CDM specifications in your kitchen.

- You’ll also need to sort and tag all your fish and flour using my jars/ontologies.

- By the way, my SAS cooking set is now available. You’ll need the double broiler to participate in the cooking club....you mean your non-profit doesn’t qualify for an academic discount?


\textsuperscript{2} Photo Credit: Mike Mozart https://www.flickr.com/photos/jeepersmedia/16058301570
How to you keep order in your kitchen? Do you need an oven for every initiative?

Top Photo Credit: Mark Nobil  https://www.flickr.com/photos/knobil/4521285655
Remember: Central CTSA Informatics Aim: Create a data “fishing” platform?

📍 Get a License: Develop business agreements, policies, data use agreements and oversight.

📍 Get a Fishing Rod and Bass Boat:
Implement open source NIH funded (i.e. i2b2 https://www.i2b2.org/) initiatives for accessing data.

📍 Know what your catching: Transform data into information using the NLM UMLS Metathesaurus as our vocabulary source.
- Secondary goal; mostly irrelevant at one site?
- !!!This is now important!!!

📍 Stock Different Tasty Fish: link clinical data sources to enhance their research utility.
PCORnet Common Data Model v3.0

**DEMOGRAPHIC** v1.0
Demographics record the direct attributes of individual patients.

**ENROLLMENT** v1.0
Enrollment is a concept that defines a period of time during which a person is expected to have complete data capture. This concept is often insurance-based, but other methods of defining enrollment are possible.

**ENCOUNTER** v1.0
Encounters are interactions between patients and providers within the context of healthcare delivery.

**DIAGNOSIS** v1.0
Diagnosis codes indicate the results of diagnostic processes and medical coding within healthcare delivery. Data in this table are expected to be from healthcare-mediated processes and reimbursement drivers.

**PROCEDURES** v1.0
Procedure codes indicate the discreet medical interventions and diagnostic testing, such as surgical procedures and lab orders, delivered within a healthcare context.

**VITAL** v1.0
Vital signs (such as height, weight, and blood pressure) directly measure an individual’s current state of attributes.

**LAB_RESULT_CM** v2.0
Laboratory result Common Measures (CM) use specific types of quantitative and qualitative measurements from blood and other body specimens. The common measures are defined in the same way across all PCORnet networks, but this table can also include other types of lab results.

**DISPENSING** v2.0
Outpatient pharmacy dispensing, such as prescriptions filled through a neighborhood pharmacy with a claim paid by an insurer. Outpatient dispensing may not be directly captured within healthcare systems.

**PRESCRIBING** v3.0
Provider orders for medication dispensing and/or administration. These orders may take place in any setting, including the inpatient or outpatient basis.

**CONDITION** v2.0
A condition represents a patient’s diagnosed and self-reported health conditions and diseases. The patient’s medical history and current state may both be represented.

**PCORNET_TRIAL** v3.0
Patients who are enrolled in PCORnet clinical trials.

**PRO_CM** v2.0
Patient-Reported Outcome (PRO) Common Measures (CM) are standardized measures that are defined in the same way across all PCORnet networks. Each measure is recorded at the individual item level: an individual question/statement, paired with its standardized response options.

**DEATH** v3.0
Reported mortality information for patients.

**DEATH_CAUSE** v3.0
The individual causes associated with a reported death.

**HARVEST** v3.0
Attributes associated with the specific PCORnet datamart implementation, including data refreshes.
GPC Ontologies -> PCORnet Common Data Model (CDM)

**GPC Progress**
- All sites have counts on Babel
  - Study population – 8.4M
  - [https://babel.gpcnetwork.org](https://babel.gpcnetwork.org)
- Annotated data dictionaries have been completed by all sites
- Software has been developed that executes against a GPC i2b2 ontology (containing CDM and additional domains such as labs, meds) to populate the CDM

**Examples on the road to data standardization**
- Specification for CDM is billing diagnosis (vs. clinical encounter diagnosis); we can supply both
- Enrollment; may have “micro” enrollments per condition (eg. Admission; Cancer); defining encounters also challenging
From Observation to Intervention
Aspirin Dosing: A Patient-Centric Trial Assessing Benefits and Long-term Effectiveness (ADAPTABLE) Trial

PCORnet’s First Pragmatic Clinical Trial

This is where real action is based on Duke’s sustainability projections.

Which medical centers/health systems can execute a new trial paradigm?

Notice also who doesn’t participate
ADAPTABLE (Aspirin Dosing: A Patient-centric Trial Assessing Benefits and Long-Term Effectiveness)

**Aim:** to compare the effectiveness and safety of two doses of aspirin (81 mg and 325 mg) in 20,000 high-risk patients with atherosclerotic cardiovascular disease (ASCVD)

**Enrollment as of 02/20/2017: >2200 across 27 sites**

**Cost per enrolled patient has come down by an order of magnitude**
- ADAPTABLE = $850
- 3 Simple, NIH Pragmatic Trials: $2,260 to $13,269
- Industry Trials: $8,500

**Key Innovation**
- Approach/platform for recruiting patients into clinical trials efficiently

**Who’s Involved**
- Researchers from 8 PCORnet CDRNs and 35 Health Systems
- Adaptors Team (8 patients representing each CDRN)
- Health eHeart Alliance PPRN support the Adaptors team
## Site Approach & Enrollment Update (4/24)

<table>
<thead>
<tr>
<th>CDRN</th>
<th>Site</th>
<th>Total Number Eligible</th>
<th>Total Number Approached</th>
<th>% of Eligible Approached</th>
<th>Golden Tickets Entered</th>
<th>% Golden Tickets entered per Approached</th>
<th>Total Enrolled</th>
<th># Non-internet Enrolled</th>
<th>% Enrolled Per Golden Ticket Entered</th>
<th>% Enrolled Per Approached</th>
<th>% Enrolled Per Approached</th>
<th>Enrolled last week</th>
</tr>
</thead>
<tbody>
<tr>
<td>MidSouth</td>
<td>Vanderbilt</td>
<td>18,440</td>
<td>6,213</td>
<td>34%</td>
<td>863</td>
<td>14%</td>
<td>459</td>
<td>12</td>
<td>7%</td>
<td>53%</td>
<td>33%</td>
<td>26</td>
</tr>
<tr>
<td>PaTH</td>
<td>UPMC</td>
<td>13,879</td>
<td>5,220</td>
<td>38%</td>
<td>756</td>
<td>14%</td>
<td>252</td>
<td>0</td>
<td>5%</td>
<td>33%</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Mid-South</td>
<td>Duke</td>
<td>19,902</td>
<td>738</td>
<td>4%</td>
<td>340</td>
<td>46%</td>
<td>241</td>
<td>33</td>
<td>33%</td>
<td>71%</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>REACHnet</td>
<td>Ochsner</td>
<td>11,832</td>
<td>6,096</td>
<td>52%</td>
<td>560</td>
<td>9%</td>
<td>213</td>
<td>37</td>
<td>3%</td>
<td>38%</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>OneFlorida</td>
<td>U of Florida</td>
<td>29,738</td>
<td>7,065</td>
<td>24%</td>
<td>163</td>
<td>2%</td>
<td>147</td>
<td>26</td>
<td>2%</td>
<td>90%</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>NYC-CDRN</td>
<td>Montefiore</td>
<td>47,383</td>
<td>1,713</td>
<td>4%</td>
<td>147</td>
<td>9%</td>
<td>121</td>
<td>48</td>
<td>7%</td>
<td>82%</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>GPC</td>
<td>KUMC</td>
<td>12,936</td>
<td>3,916</td>
<td>30%</td>
<td>273</td>
<td>7%</td>
<td>119</td>
<td>0</td>
<td>3%</td>
<td>44%</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>GPC</td>
<td>Iowa</td>
<td>14,512</td>
<td>5,274</td>
<td>36%</td>
<td>228</td>
<td>4%</td>
<td>112</td>
<td>6</td>
<td>2%</td>
<td>49%</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>PaTH</td>
<td>Penn St</td>
<td>2,708</td>
<td>2,647</td>
<td>98%</td>
<td>284</td>
<td>11%</td>
<td>100</td>
<td>0</td>
<td>4%</td>
<td>35%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PaTH</td>
<td>Utah</td>
<td>5,219</td>
<td>3,430</td>
<td>66%</td>
<td>197</td>
<td>6%</td>
<td>85</td>
<td>8</td>
<td>2%</td>
<td>43%</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>pScanner</td>
<td>UCLA</td>
<td>15,669</td>
<td>5,600</td>
<td>36%</td>
<td>134</td>
<td>2%</td>
<td>62</td>
<td>1</td>
<td>1%</td>
<td>46%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CAPriCORN</td>
<td>Northwestern</td>
<td>15,234</td>
<td>3,593</td>
<td>24%</td>
<td>114</td>
<td>3%</td>
<td>55</td>
<td>1</td>
<td>2%</td>
<td>48%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GPC</td>
<td>Marshfield Clinic</td>
<td>8,277</td>
<td>3,730</td>
<td>45%</td>
<td>108</td>
<td>3%</td>
<td>53</td>
<td>0</td>
<td>1%</td>
<td>49%</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>CAPriCORN</td>
<td>U of Chicago</td>
<td>5,446</td>
<td>543</td>
<td>10%</td>
<td>52</td>
<td>10%</td>
<td>51</td>
<td>34</td>
<td>9%</td>
<td>98%</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>GPC</td>
<td>MCW</td>
<td>5,731</td>
<td>2,741</td>
<td>48%</td>
<td>100</td>
<td>4%</td>
<td>38</td>
<td>0</td>
<td>1%</td>
<td>38%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>REACHnet</td>
<td>BSW</td>
<td>2,081</td>
<td>1,858</td>
<td>89%</td>
<td>84</td>
<td>5%</td>
<td>34</td>
<td>5</td>
<td>2%</td>
<td>40%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NYC-CDRN</td>
<td>NYU</td>
<td>49,254</td>
<td>1,126</td>
<td>2%</td>
<td>134</td>
<td>12%</td>
<td>30</td>
<td>1</td>
<td>3%</td>
<td>22%</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>NYC-CDRN</td>
<td>Weill Cornell</td>
<td>17,253</td>
<td>853</td>
<td>5%</td>
<td>94</td>
<td>11%</td>
<td>24</td>
<td>0</td>
<td>3%</td>
<td>26%</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>PaTH</td>
<td>Temple</td>
<td>4,101</td>
<td>2,179</td>
<td>53%</td>
<td>69</td>
<td>3%</td>
<td>17</td>
<td>4</td>
<td>1%</td>
<td>25%</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>CAPriCORN</td>
<td>Rush</td>
<td>8,826</td>
<td>960</td>
<td>11%</td>
<td>32</td>
<td>3%</td>
<td>14</td>
<td>0</td>
<td>1%</td>
<td>44%</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>GPC</td>
<td>UTSW</td>
<td>2,459</td>
<td>495</td>
<td>20%</td>
<td>24</td>
<td>5%</td>
<td>15</td>
<td>0</td>
<td>3%</td>
<td>63%</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>NYC-CDRN</td>
<td>Mt Sinai</td>
<td>41,876</td>
<td>261</td>
<td>1%</td>
<td>35</td>
<td>13%</td>
<td>12</td>
<td>3</td>
<td>5%</td>
<td>34%</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GPC</td>
<td>Missouri</td>
<td>1,004</td>
<td>617</td>
<td>61%</td>
<td>28</td>
<td>5%</td>
<td>10</td>
<td>0</td>
<td>2%</td>
<td>36%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mid-South</td>
<td>UNC</td>
<td>5,204</td>
<td>145</td>
<td>3%</td>
<td>21</td>
<td>14%</td>
<td>7</td>
<td>0</td>
<td>5%</td>
<td>33%</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>REACHnet</td>
<td>Tulane</td>
<td>771</td>
<td>63</td>
<td>8%</td>
<td>6</td>
<td>10%</td>
<td>3</td>
<td>2</td>
<td>5%</td>
<td>50%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GPC</td>
<td>Nebraska</td>
<td>3,200</td>
<td>102</td>
<td>3%</td>
<td>8</td>
<td>3%</td>
<td>3</td>
<td>0</td>
<td>3%</td>
<td>38%</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PaTH</td>
<td>Johns Hopkins</td>
<td>23,935</td>
<td>5</td>
<td>0%</td>
<td>1</td>
<td>20%</td>
<td>0</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td><strong>362,935</strong></td>
<td><strong>67,183</strong></td>
<td><strong>18%</strong></td>
<td><strong>4,855</strong></td>
<td><strong>7%</strong></td>
<td><strong>2277</strong></td>
<td><strong>221</strong></td>
<td><strong>3%</strong></td>
<td><strong>47%</strong></td>
<td><strong>117</strong></td>
<td><strong>117</strong></td>
</tr>
</tbody>
</table>
ADAPTABLE Enrollment

Cumulative Projected vs. Cumulative Actual Enrollment from Year 1 to Year 24.
Part 3, How does a medical center (KUMC) and the GPC (peer AMCs) fit into the evolving national landscape?

What is GPC? sites in all three national demonstration projects
  • Bariatric surgery obesity study (8 sites; 1 dropping)
  • Pediatric antibiotic obesity study (5 sites; 1 dropping)
  • ADAPTABLE Aspirin Trial (9 sites; 7 recruiting)
    ▪ leading Cancer Collaborative Research Group
    ▪ Betsy Chrichilles/Iowa and Kaiser Permanente/Portal CDRN
    ▪ responsive to National Common Data Model queries
  ▪ GROUSE claims/EMR repository

GPC Pilots

People Centered Research Foundation and speculation regarding national landscape

GPC as a data fitness camp for academic medical centers to participate in national data intensive research
Observational Obesity Studies: Both Address Controversial Topics with Largest N to Date

**Bariatric Study: what is best surgical approach?**
- 11 Participating CDRNs
- 3 Participating PPRNs
- 48 Institutions
- N=65,000 People (1,000 of whom are adolescents)

**Antibiotics given to young children: do they increase risk for obesity?**
- 10 Participating CDRNs
- 4 Participating PPRNs
- 41 Institutions
- N=650,000 children
PCORnet Pre-Research Queries

PCORnet enables large-scale data analysis across a diverse US population

N ~ 110 million patients

Same statistical code sent to Networks, which execute it against their analysis-ready data marts

Query 1: Time Required for Networks to Provide Results

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Run Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>&lt;1 minute</td>
</tr>
<tr>
<td>Maximum</td>
<td>5.5 hours</td>
</tr>
<tr>
<td>Mean</td>
<td>61 minutes</td>
</tr>
<tr>
<td>Median</td>
<td>34.5 minutes</td>
</tr>
</tbody>
</table>

EXEMPLARY
E.g. August 2016 Count Table by age, sex, race and for 10 conditions (data provided in 10.25.16 materials)
Funder query for CT for calcium scoring for prevention of CV disease (data provided in 10.25.16 materials)
GPC Query responsiveness to national office

GPC Status Report

Median Query Fulfillment Response Time: SAS vs. MDQ

<table>
<thead>
<tr>
<th></th>
<th>SAS</th>
<th>MDQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th># of Partners Who Responded</th>
<th># of Partners Who Received Query</th>
<th>GPC</th>
<th>All CDNs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>9</td>
<td>9</td>
<td></td>
<td>14.0</td>
</tr>
<tr>
<td>WCT</td>
<td>9</td>
<td>9</td>
<td></td>
<td>14.0</td>
</tr>
<tr>
<td>FD1</td>
<td>9</td>
<td>9</td>
<td></td>
<td>13.0</td>
</tr>
<tr>
<td>MDQ1</td>
<td>7</td>
<td>7</td>
<td></td>
<td>11.0</td>
</tr>
<tr>
<td>MDQ2</td>
<td>7</td>
<td>7</td>
<td></td>
<td>11.0</td>
</tr>
<tr>
<td>FD2</td>
<td>9</td>
<td>10</td>
<td></td>
<td>8.0</td>
</tr>
<tr>
<td>FD3MDQ1</td>
<td>6</td>
<td>6</td>
<td></td>
<td>11.5</td>
</tr>
<tr>
<td>FD3MDQ2</td>
<td>6</td>
<td>6</td>
<td></td>
<td>11.5</td>
</tr>
<tr>
<td>FD3MDQ3</td>
<td>6</td>
<td>6</td>
<td></td>
<td>11.5</td>
</tr>
<tr>
<td>FD3MDQ4</td>
<td>6</td>
<td>6</td>
<td></td>
<td>11.5</td>
</tr>
<tr>
<td>FD3MDQ5</td>
<td>6</td>
<td>6</td>
<td></td>
<td>11.5</td>
</tr>
<tr>
<td>FD4MDQ1</td>
<td>7</td>
<td>7</td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td>FD4MDQ2</td>
<td>7</td>
<td>7</td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td>FD3</td>
<td>9</td>
<td>11</td>
<td></td>
<td>7.0</td>
</tr>
</tbody>
</table>

Data Sharing Agreement (DSA) V.1.0*

<table>
<thead>
<tr>
<th></th>
<th>Signed DSA</th>
<th>Total # of Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPC</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

SMART IRB*

<table>
<thead>
<tr>
<th></th>
<th>Executed SMART IRB</th>
<th>Total # of sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPC</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>

*Have Not Executed SMART IRB
** University of Missouri (MU)**
** = Initiated process

Query Fulfillment Status:
FD1: Children’s Mercy (rejected)
FD2: Children’s Mercy (rejected)
FD3: Children’s Mercy (rejected), UI HealthCare & UTSW (no response)

Network Partners:
- Children’s Mercy
- IU
- KUMC
- Marshfield
- MCW
- MU
- UI HealthCare
- UMN
- UNMC
- UTHSCSA
- UTSW
- UW-Madison

*Denominators vary by metric
GPC Regulatory responsiveness to national

PCORnet Status Update for CDRNs

**RATIONALE 1**
SMART IRB

Percent of CDRN sites who have executed SMART IRB

**RATIONALE 2**
Data Sharing Agreement (DSA) V.1.0

Number of Sites with Fully-Executed DSAs

**RATIONALE 3**
Query Fulfillment

Median Network Query Response Time*

---

**RATIONALE 1**: NIH will require a single IRB for sites participating in federally-funded multi-site trials beginning in September 25, 2017. PCORnet is partnering with NCATS to support sites to sign on to the SMART IRB single IRB model and obtain a fully executed jinder agreement.

**RATIONALE 2**: The Data Sharing Agreement (DSA) defines the standard terms to which the PCORnet Coordinating Center (Harvard Pilgrim and Duke) will adhere when information/data are sent from the Network Partners to CDRN the Coordinating Center. A total of 97.5% of eligible CDRNs and PPRNs (sites) have signed an initial "V.1.0" Data Sharing Agreement. Version 2.0 (V.2.0) of the DSA is anticipated for release in March of 2017 and will be made available to the Network Partners to sign.

**RATIONALE 3**: Query Fulfillment is a complex, iterative process requiring knowledge of PCORnet Network Partner data, the common data model, existing PCORnet SAS programs or tools (e.g., PMP, MDQ) or requests for SAS program development. SAS programmed queries and menu-driven queries (MDQ) have been released to the Network Partners for query fulfillment. Quickly responding to these query requests by providing output data that contributes to a final report is essential to the success of PCORnet. Network Partners only receive queries that are relevant to them to respond to (e.g., because of the population or topic of the query).
GPC can integrate, query and ship data

- Notes integration
- Supporting CTSA i2b2/REDcap and sending data to investigators
- National intersection to support distributed research network
  - Common Data Model
  - Data Quality assessment
- GROUSE: claims integration across states for complete data assessment and comparison with state population
GPC Reusable Observable Study Environment (GROUSE) generally tertiary care sites for their regions…how biased are we…what do we not know?

- GPC centers in eight states
  - less than 20% US population
- We want to measure the gain by integrating claims for our 3 cohorts and serve as a basis for future projects that want to reuse this resource (with CMS approval).
- Originally would have desired distributed approach but are creating a central de-identified resource due to
  - CMS position when we developed plan in 2015
  - Need consistent data management plans at all sites
  - Timing to obligate money
  - Cost to manage distribution across all sites
GROUSE: Claims Files Requested to Support Aims

- 3 years of data for Medicare (2011-2013) and one year Medicaid (2011) which was the freshest available when we started this process for phase 1.
  - Cost was ~$190,000 for our eight states
  - Purchasing refreshed Medicare (2014,2015) and Medicaid (2012-2013)

<table>
<thead>
<tr>
<th>Study objective/measure</th>
<th>Data variable(s)</th>
<th>CMS data file(s) needed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment patterns</td>
<td>Diagnosis, procedures, medications</td>
<td>Outpatient, Carrier, Part D Event &amp; Characteristic, MedPAR, MAX Inpatient, MAX Prescription, MAX Other Therapies</td>
</tr>
<tr>
<td>Outcomes</td>
<td>LTC placement, hospitalization</td>
<td>Outpatient, MedPAR, MAX Inpatient, MAX Long Term Care</td>
</tr>
<tr>
<td>Distribution of conditions across regions</td>
<td>Diagnosis, demographics</td>
<td>Outpatient, Carrier, Master Beneficiary Summary File (A/B/D) Segment, MedPAR, MAX Person Summary, MAX Inpatient</td>
</tr>
<tr>
<td>Cohort characterization</td>
<td>Diagnosis, procedures, medications</td>
<td>Outpatient, Carrier, Part D Event &amp; Characteristic, MedPAR, MAX Inpatient, MAX Prescription, MAX Other Therapies</td>
</tr>
</tbody>
</table>
**GROUSE: Current work**

- Received and staged claims (~1.2 terabytes)
  - 3.8 billion rows, 19 million beneficiaries,
  - Ex: hospice claims for ALS patients:

- Sites generated finder files. Crosswalk received from GDIT March!
  - Have BMI to claims linkage on ~150,000 Medicare/ 50,000 Medicaid beneficiaries at KUMC alone.

- Prioritizing data formats and environments for investigators.
  - De-id raw CMS files
  - CDM ETL of CMS files
  - i2b2 data with CMS data
  - SAS config
  - R environment

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>736</td>
<td>Texas</td>
</tr>
<tr>
<td>294</td>
<td>Wisconsin</td>
</tr>
<tr>
<td>286</td>
<td>Missouri</td>
</tr>
<tr>
<td>280</td>
<td>Indiana</td>
</tr>
<tr>
<td>249</td>
<td>Minnesota</td>
</tr>
<tr>
<td>169</td>
<td>Iowa</td>
</tr>
<tr>
<td>131</td>
<td>Kansas</td>
</tr>
<tr>
<td>84</td>
<td>Nebraska</td>
</tr>
</tbody>
</table>
Pilots: Use GPC to generalize and increase GPC investigators’ competitiveness.

Capabilities and Pilot Categories
- Use GROUSE (Claims + EMR + Registry + Socio-demographics) for Breast Cancer, ALS, or Height/Weight implication on health or feasibility queries. (~<$5000)
- Multi-site de-identified or limited dataset as observational study or prelim data for trial (~$10,000)
- Survey or PRO deployment across sites via REDCap like our initial cohorts in phase 1 ($20,000+)

Support models:
- Have some modest explicit funding from GPC ($~$60,000)
- In kind support and matching funds from GPC informatics and CTSAs
  - match where pilots that are led by their investigators or co-I

Timeline: Announced April 10th, LoI May 31st, Due July 21st, Award August 31st
- Work with CTSA programs to provide review
- Invite awardees to our annual Fall event (GPC LEC) in October
  (http://gpcnetwork.org/?q=events)
  - GPC would fund awardees to attend and have sites and patients engage to develop stronger projects and national proposals
The PCORI contracts to CDRNs were for 4.5 years to then be sustained by network’s research.

There was an election in November. Would PCORI be reauthorized?
- PCORnet was built with PCORI funding as a major sustaining component
- FDA as sustaining partner, enhanced by Rob Califf’s role as commissioner

The People-Centered Research Foundation (PCRF) is committed to accelerating people-driven research that is faster, more user friendly for patients and providers, and less costly. Our mission is: To engage patients, families, research participants, clinicians, scientists, and health system leaders in the design, conduct, dissemination, and implementation of research and analysis that leads to improvements in the health and well-being of individuals and populations and the performance of health care delivery systems.

The PCRF was established to sustain and expand a national network for clinical research that originated with funding from the Patient-Centered Outcomes Research Institute (PCORI) and studies conducted by The National Patient-Centered Clinical Research Network (PCORnet). PCORnet has led the way to transform the culture of clinical research by creating a system driven by the shared needs of people and their families, communities, clinicians, and health system leaders. We are ensuring PCORnet’s vision and unprecedented work continue.

Robert Califf, former FDA Commissioner and Duke faculty, Chair
- PCORI board voted > $25 million initial start up funds.
Historical model of clinical research: Many recruitment sites and a coordinating center

- Hub & spoke model
- Top-down decision-making
- Sites operated independently

Interoperable Networks Share Sites and Data “reciprocity”

Each organization can participate in multiple networks
Each network controls its governance and coordination
Networks share infrastructure, data curation, analytics, lessons, security, software development
Other potential partners: disease or treatment-specific networks;

To date: PCORnet has mostly followed hierarchical model given contract structure

GPC: intermediary, member governed collaborative

- Helps member improve their data infrastructure to participate in national research
- Contracting for PCORnet/NewCo or sponsor
- Governs peer to peer data sharing to complement SMART IRB
- Provides a forum for members to share technology and be accountable to one another for data quality and capability
- Can consolidate data assets if needed such as CMS claims via GROUSE
GPC Future Directions and Discussion with our CTSA Leaders

Current GPC is managed as a contract so our data sharing language references the master contract.

- As we move beyond the 18 months, it will be more cost effective and useful to modify the language to add interested partners: sharing management costs and enabling larger studies or likelihood our investigators will find their desired peers covered by the GPC agreements and data interoperability. (1.1 change)

PCORnet’s pursing a new national nonprofit that would seek to contract with networks.

- Seeking PCORI Board investment to covering funding shortfall (~$50 million) until estimated breakeven in 5 years to sustain ~7 networks & 50-70 nodes

Long term thinking from the GPC Governing Council is to continue GPC as a collaborative (not a separate entity) governed by participating/performing sites via a charter (2.0 change)

- Sustain via membership dues like Kuali or HSCRN (formerly known as HMO Research Network)
- Much less ambitious but “Ramen sustainable” ~$300-500k/year ($15-30k/member)

We think GPC can be responsive using our current model or as a collaborative guided by a new charter.

- Will schedule a Summer call in to share further details on PCORnet’s post phase 2 contract developments.

More has happened than ever before to mobilize data from the point of care for research

- Data and technology are fluid and we’re in a dynamic time: leaders will master interventional informatics and engagement for both research and care delivery
Questions?