The Greater Plains Collaborative: a PCORNet Clinical Data Research Network’s Strategies for Creating an Interoperable Architecture

“From Fishing to Attracting Chicks”

Russ Waitman, PhD
Associate Professor, Internal Medicine
Director, Medical Informatics
Assistant Vice Chancellor, Enterprise Analytics
University of Kansas Medical Center
Kansas City, Kansas

This project is supported in part by NIH grant UL1TR000001 and PCORI Contract CDRN-1306-04631
Two Analogies: Local to National

• “Hooking” People on Informatics
• “Luring” Researchers for collaboration at a national level
• Demo of Babel
Sample CTSA Approach from KUMC:
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** – *fish for data* from the hospital/clinic

• **Biweekly** Frontiers Clinical Informatics **Clinics**
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
Why so many fish?

**Current Goal:** Build Hatchery, Manage the Fishery
Second Goal: If you need help fishing, get a guide

Photo Credit: HuntFishGuide.com
http://www.flickr.com/photos/huntfishguide/5883317106/
Prepare and Analyze Data

Photo Credit: S. Klathill
http://www.flickr.com/photos/sklathill/505464990/
Our shared goal: a tasty publication

Photo Credit: Steve Velo
http://www.flickr.com/photos/juniorvelo/259888572/
Fishing in a Learning Healthcare System

• Evidence Generation: Trials post Electronic Health Records (EHR)
  – 2011 Report: *Digital Infrastructure for the Learning Health System: The Foundation for Continuous Improvement in Health and Health Care*
    • “*Information Technology (IT) serves as the functional engine for the continuous learning system*”

• Effective Biomedical Research Organizations will integrate Medical Informatics and Clinical Research Informatics capabilities with Operational Clinical Systems of the Healthcare Systems
  – EHR Standardization and collection of data becomes critical to the cost effective research capacity of the University
  – EHRs (Cerner/Epic), clinical research systems like (REDCap/Oncore/Velos), and ultimately translational bioinformatics systems will need to be managed in an integrated environment and NOT siloed.
The “Greater Plains Collaborative”

- KS, the University of Kansas Medical Center (KUMC)
- MO, Children’s Mercy Hospital
- IA, University of Iowa Healthcare
- 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.
- Selected in July to submit full proposal in September, award in December, funded March 6, 2014
  - $7 million total costs over 18 months
Greater Plains Collaborative Objective

• Lek: gather in the the Spring on a Booming Ground to attract female Greater Prairie Chickens
• If you dance by yourself, you’re not attracting researchers interested in generalizable results
• GPC: CTSAs create ideal habitat for clinical researchers to come and study our state’s populations and develop methods to improve our communities health outcomes
• Data, IRB, and governance in place so we can enable Comparative Effectiveness Research trials and Patient Reported Outcome collection

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
Support Biospecimen Requests
Common CTSA Informatics Aim: Create a data warehouse such as HERON at KUMC

- **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 and standards as our vocabulary source.

- **Stock Different Tasty Fish:** extract clinical data sources (Epic, Cerner, Cattails) to integration platform to enhance their research utility.
• GPC: initially characterizing three cohorts
  • Use honest brokers for preliminary counts and data extraction
  • Focus on Data Sharing Agreement to support observational studies with de-identified data
    • GPC IRB Initial Consensus: IRB protocol not required for fully de-identified data
  • Limited datasets to monitor prospective Comparative Effectiveness trials that have IRB approval

Single sign-on using your email username

Real-time check for current human subjects training
Governance:
1. IRB protocol for honest broker and data repository
2. IRB approval for prospective trials
3. HealthSystem oversight for de-identified data requests
4. CTSA assistance and oversight for contacting patients
5. HealthSystem CMIO/CIO oversight of EMR modifications and components

Concepts paths shared for interoperability efforts and network marketing

After approval, site honest brokers run extract scripts from site i2b2 -> GPC i2b2 or REDCap data store
Anticipate a Two Phase Approach to Data Standardization and Interoperability:

1. **Post ETL Concept Mapping by Informatics**
   - **MU2 Alignment by EMR Build Team**
     - Local index term: 100%
     - Inventory term: 99%
     - FDB for Decision Support: 90%
     - Name: 100%
     - **Interoperable Standard: 94%**
   - **MU2+i2b2 alignment**
     - Local index term: 100%
     - Inventory term: 99%
     - FDB for Decision Support: 90%
     - Name: 100%
     - **Interoperable Standard: 98%**
   - **Cohort Characterization Using Interoperable Standards**
     - (months 13-18)

2. **EMR based Meaningful Use activity**
   - **Initial Cohort Characterization Activities**
     - (months 6-14)
   - **Cross site i2b2-based alignment by GPC**
     - Local index term: 100%
     - Inventory term: 95%
     - FDB for Decision Support: 75%
     - Name: 100%
     - **Interoperable Standard: 95%**
   - **PROM Mockups in REDCap**
     - - align with MU

**Concept Example: Medications**
- Local index term: Medication_ID 452
- Inventory term: NDC Code 54868-3105-0
- FDB for Decision Support: gcn_seqno 9001
- Name: AMOXICILLIN 250 MG PO CHEW
- Interoperable Standard: RXOUI 370577
- Standard Name: Amoxicillin Chewable Tablet
The “Greater Plains Collaborative”
Support for interoperation

• Research data integration and management tooling: i2b2
• Information model: Star schema
• Domain ontology/code sets:
  – Demographics, Clinical findings/biometrics, Lab findings, Radiology findings, Diagnoses, Allergies, Procedures, Orders - procedure/medications, Medications/pharmaceuticals administered, Registry data
• Value sets for coded data
• Phase 1: modify concept mapping code at sites
  – KUMC examples: _concepts.sql at https://informatics.kumc.edu/work/browser/heron_load
The “Greater Plains Collaborative”
Epic EHR Sites (8 of 10): Clarity Data Resources

- **“Fish”** through Clarity Data Dictionary, site workflows and Epic build to identify the datasets
- Map Epic EHR data to vocabulary standards in synchronization with Meaningful Use requirements
- Collaborate with Clarity data management team to assure that necessary extract tables are populated
- Manage extract cycle to assure that timely standardized data is delivered to i2b2
- Employ i2b2 integrated data sets to support quality assurance and research data management
The “Greater Plains Collaborative”
Epic EHR Sites (8 of 10): Clarity Data Resources

- “Fish” through Clarity Data Dictionary, site workflows and Epic build to identify the datasets
- Map Epic EHR data to vocabulary standards in synchronization with Meaningful Use requirements
- Collaborate with Clarity data management team to assure that

The “Greater Plains Collaborative” Epic EHR Sites (8 of 10): Clarity Data Resources

- “Fish” through Clarity Data Dictionary, site workflows and Epic build to identify the datasets
- Map Epic EHR data to vocabulary standards in synchronization with Meaningful Use requirements
- Collaborate with Clarity data management team to assure that

The “Greater Plains Collaborative” Epic EHR Sites (8 of 10): Clarity Data Resources

- “Fish” through Clarity Data Dictionary, site workflows and Epic build to identify the datasets
- Map Epic EHR data to vocabulary standards in synchronization with Meaningful Use requirements
- Collaborate with Clarity data management team to assure that

The “Greater Plains Collaborative” Epic EHR Sites (8 of 10): Clarity Data Resources

- “Fish” through Clarity Data Dictionary, site workflows and Epic build to identify the datasets
- Map Epic EHR data to vocabulary standards in synchronization with Meaningful Use requirements
- Collaborate with Clarity data management team to assure that

The “Greater Plains Collaborative” Epic EHR Sites (8 of 10): Clarity Data Resources

- “Fish” through Clarity Data Dictionary, site workflows and Epic build to identify the datasets
- Map Epic EHR data to vocabulary standards in synchronization with Meaningful Use requirements
- Collaborate with Clarity data management team to assure that

The “Greater Plains Collaborative” Epic EHR Sites (8 of 10): Clarity Data Resources

- “Fish” through Clarity Data Dictionary, site workflows and Epic build to identify the datasets
- Map Epic EHR data to vocabulary standards in synchronization with Meaningful Use requirements
- Collaborate with Clarity data management team to assure that

The “Greater Plains Collaborative” Epic EHR Sites (8 of 10): Clarity Data Resources

- “Fish” through Clarity Data Dictionary, site workflows and Epic build to identify the datasets
- Map Epic EHR data to vocabulary standards in synchronization with Meaningful Use requirements
- Collaborate with Clarity data management team to assure that

The “Greater Plains Collaborative” Epic EHR Sites (8 of 10): Clarity Data Resources

- “Fish” through Clarity Data Dictionary, site workflows and Epic build to identify the datasets
- Map Epic EHR data to vocabulary standards in synchronization with Meaningful Use requirements
- Collaborate with Clarity data management team to assure that

The “Greater Plains Collaborative” Epic EHR Sites (8 of 10): Clarity Data Resources

- “Fish” through Clarity Data Dictionary, site workflows and Epic build to identify the datasets
- Map Epic EHR data to vocabulary standards in synchronization with Meaningful Use requirements
- Collaborate with Clarity data management team to assure that

The “Greater Plains Collaborative” Epic EHR Sites (8 of 10): Clarity Data Resources

- “Fish” through Clarity Data Dictionary, site workflows and Epic build to identify the datasets
- Map Epic EHR data to vocabulary standards in synchronization with Meaningful Use requirements
- Collaborate with Clarity data management team to assure that

The “Greater Plains Collaborative” Epic EHR Sites (8 of 10): Clarity Data Resources

- “Fish” through Clarity Data Dictionary, site workflows and Epic build to identify the datasets
- Map Epic EHR data to vocabulary standards in synchronization with Meaningful Use requirements
- Collaborate with Clarity data management team to assure that
The “Greater Plains Collaborative”
Epic EHR Sites (8 of 10): Clarity Data Resources

- “Fish” through Clarity Data Dictionary, site workflows and Epic build to identify the datasets
- Map Epic EHR data to vocabulary standards in synchronization with Meaningful Use requirements

Meaningful Use Terminology Mapping Guide

Legends:
- Black: Items are current site processes/systems.
- Green: Items are data sources which might be piloted at each site, but not deployed across the network.
- Red: Items are new components deployed at each site across the network.
- Blue: Items are components deployed centrally.
- Purple: Lines show the feedback processes to configure sites for PROM, CER, and coordinating amongst biospecimen repositories.
The “Greater Plains Collaborative” Meaningful Use Vocabulary Standards

- Demographics (80% stage 2): HL7/OMB code set
- Family history, past medical history, smoking status, clinical observations: SNOMED CT
- Problem list/diagnoses (80% of patients): SNOMED CT, ICD*
- Structured lab results (55% stage 2): Lab LOINC
- E-prescribing (50% formulary check stage 2): RxNORM
- Medications: RxNORM
- Immunizations (Immunization registries): CVX, MVX
- Procedures (summary of care): CPT, HCPCS
- Documents (summary of care): LOINC
# The “Greater Plains Collaborative” ONC Stage 1&2 Vocabulary Standards

<table>
<thead>
<tr>
<th>Domain</th>
<th>Ontology/Code sets</th>
<th>Value sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td>(Clinical LOINC)</td>
<td>HL7/OMB</td>
</tr>
<tr>
<td>Diagnoses</td>
<td>SNOMED CT; ICD-9-CM; ICD-10-CM</td>
<td></td>
</tr>
<tr>
<td>Clinical findings</td>
<td>SNOMED CT (Clinical LOINC)</td>
<td></td>
</tr>
<tr>
<td>Lab findings</td>
<td>Lab LOINC</td>
<td>SNOMED CT</td>
</tr>
<tr>
<td>Allergies</td>
<td>SNOMED CT; RXNORM</td>
<td></td>
</tr>
<tr>
<td>Procedures</td>
<td>CPT, HCPCS, SNOMED CT</td>
<td></td>
</tr>
<tr>
<td>Medication orders</td>
<td>RXNORM</td>
<td></td>
</tr>
</tbody>
</table>
## The “Greater Plains Collaborative”
### Epic EHR Sites: Vocabulary Deployment

<table>
<thead>
<tr>
<th>Domain</th>
<th>Ontology/Code sets</th>
<th>Value sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
<td>HL7/OMB</td>
</tr>
<tr>
<td>Diagnoses</td>
<td>SNOMED CT; ICD-9-CM; ICD-10-CM (IMO)</td>
<td></td>
</tr>
<tr>
<td>Clinical findings</td>
<td>SNOMED CT (Clinical LOINC)</td>
<td></td>
</tr>
<tr>
<td>Lab findings</td>
<td>Lab LOINC</td>
<td>SNOMED CT</td>
</tr>
<tr>
<td>Allergies</td>
<td>SNOMED CT; RXNORM</td>
<td></td>
</tr>
<tr>
<td>Procedures</td>
<td>CPT, HCPCS, SNOMED CT</td>
<td></td>
</tr>
<tr>
<td>Medication orders</td>
<td>RXNORM</td>
<td></td>
</tr>
</tbody>
</table>

**Available in most implementations via IMO**

**Must be mapped per Epic**

**Requires extension of Epic data model**
Current Approach: Babel, an i2b2 Terminology Repository

- Babel.gpcnetwork.org
- Amazon Web Services (established HIPAA Business Associates Agreement); subsequent infrastructure filesystem encryption
- Concept paths only; no underlying derived data
- Statistics scripts to calculate facts and patients at each node in ontologies
- University of Kansas
- Marshfield Clinic
- Medical College of Wisconsin + other sites
GPC common hierarchies for cross site queries
PCORNet Common Data Model Tables represented as i2b2 “Folders” and modifiers.

Calculate Statistics of what portion of data within the GPC maps to the PCORNet data model versus GPC and site level mappings.

Transpose data from EAV into SAS files compliant with PCORNet Common Data Model: house at the GPC level.

Collaborate with other i2b2 based networks on common scripts.
Relevant Epic (customers only) and GPC References

- Epic Clarity data dictionary: [https://galaxy.epic.com/Search/GetFile?url=7900%2144%21100%21212145262](https://galaxy.epic.com/Search/GetFile?url=7900%2144%21100%21212145262)
- MU stage 1&2 mapping requirements: [https://galaxy.epic.com/Search/GetFile?url=1%2168%21100%21212154008](https://galaxy.epic.com/Search/GetFile?url=1%2168%21100%21212154008)
- GPC Development: [http://informatics.gpcnetwork.org](http://informatics.gpcnetwork.org)
- GPC: [http://www.gpcnetwork.org](http://www.gpcnetwork.org)

rwaitman@kumc.edu