Greater Plains Collaborative Hackathon
PCORNet Act One

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

February 24, 2014

This project is supported in part by NIH grant UL1TR000001 and NSF Award CNS-1258315
Three Analogies, Introductions

- Part 1: “Hooking” People on Informatics
- Part 2: Clinical Integrated Data Repositories: a Cocktail Drinker’s Perspective
- Part 3: Drinking leads to....
- Team Introductions and Goals for Hackathon
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**
  – Tuesday 4-5 pm in 1028 Dykes Library.
  – Next session March 4\(^{th}\)
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/
Nightmare: looks like a nice river, but can’t catch fish

Little White Salmon River, Washington State, last Summer in July

- I’ll just enter everything in Excel....
- What if I lose or accidentally sort my spreadsheet?
- How to I let students only review de-identified data?
- Hospital/Clinic is making me use this Electronic Medical Record and I get nothing in return...
Sometimes, You’re willing to enter data/buy fish: REDCap: Research Electronic Data Capture

- **https://redcap.kumc.edu**
  - It uses the same username and password as your KUMC email.
  - Non-KUMC researchers can request an affiliate account through Frontiers CTSA office
  - Check out the training materials under videos
- **Case Report Forms**
  - (see left) registries, trials
- **Conduct Surveys**
  - Similar to “SurveyMonkey”
Vexed by a complicated protocol in a vulnerable population?

Want to go beyond compliance hurdles to sleeping well at night?

Frontiers CTSA brings you an exciting opportunity to obtain thoughtful ethics consults to further your research.

Please take a minute to fill out your request so we can direct the best ethicist to work with you on your urgent moral quandary.

First name:
* must provide value

Last Name
* must provide value

Email
* must provide value

Where are you?
* must provide value
- UMKC
- Children’s Mercy
- KCUMB
- KUMC-KC
- KUMC-W
- SLU
- KU L
- KC VA
- Other

Role
* must provide value
- Principal investigator
- Co-investigator
- Study Coordinator
- IRB member
- Other

Is your question regarding any of the common ethics topics:
- Informed Consent
- Inclusion/Exclusion criteria
- Recruitment
- Vulnerable populations

Add any additional notes or comments here

Upload your IRB protocol or a copy of your credit card here

Submit
REDCap: think Fish Tank you manage

http://www.flickr.com/photos/wiccked/185270913/lightbox/
I want to go fishing, not fill a fish tank (REDCap)
Use HERON: a managed fishery

Bonneville Hatchery: Trout, Salmon, Sturgeon, Columbia River, Oregon
Central CTSA Informatics Aim: Create a data “fishing” platform: HERON, https://heron.kumc.edu

• **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.
HERON: Getting a Fishing License

- 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
Dragging over the second condition
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

<table>
<thead>
<tr>
<th>Person_#1774_f</th>
<th>Non-hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI (Calculated): 31.070</td>
<td>80,395 patients</td>
</tr>
<tr>
<td>Diabetes mellitus: 376,016 facts; 31,763 patients</td>
<td></td>
</tr>
<tr>
<td>Frontiers Research Participant Registry: 7,654 facts; 7,654 patients</td>
<td></td>
</tr>
<tr>
<td>Glucose: 2011: 1,007,666 facts; 119,524 patients</td>
<td></td>
</tr>
<tr>
<td>HbA1c: 2013: 80,703 facts; 35,269 patients</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Person_#1809_f</th>
<th>Non-hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI (Calculated): 30.170</td>
<td>89,395 patients</td>
</tr>
<tr>
<td>Diabetes mellitus: 376,018 facts; 31,763 patients</td>
<td></td>
</tr>
<tr>
<td>Frontiers Research Participant Registry: 7,654 facts; 7,654 patients</td>
<td></td>
</tr>
<tr>
<td>Glucose: 2010: 872,682 facts; 100,192 patients</td>
<td></td>
</tr>
<tr>
<td>HbA1c: 2011: 80,703 facts; 35,269 patients</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Person_#21253_f</th>
<th>Non-hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI (Calculated): 30.170</td>
<td>89,395 patients</td>
</tr>
<tr>
<td>Diabetes mellitus: 376,018 facts; 31,763 patients</td>
<td></td>
</tr>
<tr>
<td>Frontiers Research Participant Registry: 7,654 facts; 7,654 patients</td>
<td></td>
</tr>
<tr>
<td>Glucose: 2010: 872,682 facts; 100,192 patients</td>
<td></td>
</tr>
<tr>
<td>HbA1c: 2011: 80,703 facts; 35,269 patients</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Person_#13473_f</th>
<th>Non-hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI (Calculated): 30.170</td>
<td>89,395 patients</td>
</tr>
<tr>
<td>Diabetes mellitus: 376,018 facts; 31,763 patients</td>
<td></td>
</tr>
<tr>
<td>Frontiers Research Participant Registry: 7,654 facts; 7,654 patients</td>
<td></td>
</tr>
<tr>
<td>Glucose: 2010: 872,682 facts; 100,192 patients</td>
<td></td>
</tr>
<tr>
<td>HbA1c: 2011: 80,703 facts; 35,269 patients</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Person_#5138_m</th>
<th>Non-hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI (Calculated): 30.210</td>
<td>80,395 patients</td>
</tr>
<tr>
<td>Diabetes mellitus: 376,016 facts; 31,763 patients</td>
<td></td>
</tr>
<tr>
<td>Frontiers Research Participant Registry: 7,654 facts; 7,654 patients</td>
<td></td>
</tr>
<tr>
<td>Glucose: 2010: 872,682 facts; 100,192 patients</td>
<td></td>
</tr>
<tr>
<td>HbA1c: 2011: 80,703 facts; 35,269 patients</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Person_#25806_f</th>
<th>Non-hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI (Calculated): 30.210</td>
<td>80,395 patients</td>
</tr>
<tr>
<td>Diabetes mellitus: 376,016 facts; 31,763 patients</td>
<td></td>
</tr>
<tr>
<td>Frontiers Research Participant Registry: 7,654 facts; 7,654 patients</td>
<td></td>
</tr>
<tr>
<td>Glucose: 2010: 872,682 facts; 100,192 patients</td>
<td></td>
</tr>
</tbody>
</table>
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
Part 2, Integrated Data Repositories: “A Cocktail Drinker’s Perspective”

• Kansas City: a city of contrasts as illustrated by alcohol
• Kansas: Carrie Nation, temperance society
• Missouri: “Boss” Tom Pendergast, prohibition what?
“Prohibition Never Existed” in Kansas City

http://en.wikipedia.org/wiki/History_of_the_Kansas_City_metropolitan_area#Prohibition
Echoes of these Contrasts Today:
Liquor in Gas Stations & Groceries in KCMO...

Missouri Recently identified as the #1 State for Drinking
http://business.time.com/2013/12/05/the-3-best-and-3-worst-states-in-america-for-drinking/
39th Street: next block over from KUMC/KUH
GPC Technical Mission for PCORNet
Clearly Define and Serve a Quality Drink for Research
HERON’s Data Sources, Types of Data, Gets Complex

Incorporate Clinical, Administrative, Research Datasources

<table>
<thead>
<tr>
<th>Source</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inpatient and outpatient electronic medical records (Epic)</td>
<td>Ambulatory visit details, Ordersets &amp; Decision Support used, enhance notes extract (pending)</td>
</tr>
<tr>
<td>Professional Services Billing and Scheduling (GE IDX)</td>
<td>Provider services (pending)</td>
</tr>
<tr>
<td>KUCC Biospecimen Shared Resource Samples Database</td>
<td>eSample enhance (pending)</td>
</tr>
<tr>
<td>Hospital (KUH) Tumor Registry (NAACR format)</td>
<td>enhance ontology (e.g. ER positive) (pending)</td>
</tr>
<tr>
<td>Social Security Death Master File (NIST format)</td>
<td></td>
</tr>
<tr>
<td>Technical Charges from hospital and clinics (UHC validated format)</td>
<td>Costs, clinics, ED (pending)</td>
</tr>
<tr>
<td>Research Data Capture (REDCap)</td>
<td>Replicate security model</td>
</tr>
<tr>
<td>Clinical Research Information System (Velos)</td>
<td>Clinical Trials Participation (pending)</td>
</tr>
</tbody>
</table>

HERON’s current contents

- Demographics (master patient index)
- Race/Ethnicity
- Laboratory Results
- Nursing observations/vital signs
- Clinical Diagnoses (ICD9)
- Medications (dispensed, ordered, home meds, administered)
- Physician Orders
- Procedure charges (CPT)
- Outpatient Billing diagnoses (ICD9)
- Inpatient visit/provider service
- Family/Social/Surgical Past Medical History
- Microbiology
- Allergies/Reactions
- Specimen collected
- Tumor Staging and Grade
- Diagnosis and Treatment
- Survival and Progression
- Death per Social Security Administration
- MDSRG, APDRG, LOS, Readmissions
- Technical Charge Diagnoses ICD9
- Service line, AHRQ quality and JCAHO core measures
- Triple Negative Breast Cancer Registry initial pilot completed

Status as of July 18, 2013

https://bmi-work.kumc.edu/work/wiki/HeronProjectTimeline#July2013Planning - contains current plan for next several monthly releases
“Signals” of Health: Think of each source electronic system as an EKG, blood pressure cuff, arterial line

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

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

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 app

"Health" or physiologic reserve
- Restores health
- Body’s control system + time

Healthcare Encounter "Black box" $v_t$

Provider

Observations (not recorded in EMR)
Observations about EMR use $e_1$

EMR: chart & document

Clinical Decision Support & Information Synthesis Tools

Observations about CDS effectiveness $e_2$

Actions (not recorded in EMR)

Interventions $i_t$

Ex: Laboratory, Radiology, Pathology/ exploratory procedures

Collect more data: Order Tests

Medications (orders, prescriptions)

Procedures (surgery, PT/OT, amputation)

Unregulated changes (diet)

Charges gleaned from EMR and paper $c_t$

Billing system (technical charges)

Billing system (professional charges)

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

Outcomes (not typically measured)

"Health" or physiologic reserve

Administrative billing activities collect "signals" as charges in parallel with clinical activities
Repurposing i2b2 Clinical Research Infrastructure for Inpatient Quality Improvement

- i2b2 “largely” ambulatory or population/genomics focused
- Is i2b2 version 1.6 with same financial encounter and modifiers now useful for inpatient research and precise attribution?
Initial GPC task is to set up the bar: clear bottles, clean glasses, bartender’s guide, and honest pour

- Southern Bible Belt had social clubs where the bottle belonged to members
  - Hospital, clinic, university, payor, and patient data
- Reproducible science: clear attribution wrt how data was obtained
- i2b2: an open source “clear glass” to transparently hold data
- ETL: stage bottles, clear recipes for cocktails into i2b2 glass
- Defining site/GPC oversight (HERON as basis) to oversee the pour
Reproducible Research: Modifiers, Concept Paths and sharing ETL code

- Clear attribution likely as important as data standards
- Billing, EMR, problem list, administration versus dispense?
- Load with the local terminology first; don’t miss unmapped observations.
- Map to standards secondarily in the concept space.
- Works around mapping challenges with contributing organizations; support multiple organizing hierarchies
Part 3: 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 holistically as opposed to siloed.
Patient-Centered Outcomes Research Institute
Clinical Data Research Network Announcement

- April 23, 2013
- Two types of component networks: systems-generated and patient driven networks
- Active involvement of health care systems, of clinicians and of patients
- A commitment to establishing interoperability and data sharing across networks and ready collaboration with the larger research community

What Do They Want the Network to Do?

Capabilities:

- Rich clinical data from electronic health records and other data sources stored in standardized, interoperable formats.
- Capacity to conduct both observational studies and randomized trials embedded in clinical settings
- Rigorous practices for data security and confidentiality
- Appropriate IRB and human subjects oversight
- Utility for CER, safety studies, surveillance, etiologic research, and potentially for pre-approval trials

COOPERATIVE AGREEMENT AWARD

- At least two health care systems engaged
- Willingness and capacity to work toward data standardization with other awardees
- Willingness to participate in collaborative studies with data sharing as part of a national research infrastructure

18 MONTHS LATER

- >1,000,000 patients enrolled
- Data standardized within network and with other awardee networks
- Patients, system, and clinicians engaged in governance & use
- Capable of implementing clinical trials

Joe V. Selby, MD, MPH, Executive Director ONC HITCP  June 5, 2013
The “Greater Plains Collaborative”
Awarded December! Part of pcornet.org

- **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, funding January?
- $7 million total costs over 18 months
Drinking may lead to... Dancing

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
• Get i2b2 foundation in place so we can next tackle RCT interventions and PROM collection
The “Greater Plains Collaborative”
Size, Goals, Structure

- 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
Goal: lifetime data density; data standardization and interoperability between systems and networks

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.
Introductions, what you hope to get out of Hackathon (needs may vary)

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

Environment
- "toxic sludge worker"

Social Behavioral
- "exercises by jogging 2/week"

Genetic
- " Poor warfarin metabolizer"

Diets, drugs
- "carrots, alcohol, BetaBlocker"

Illness/complaint
- "assaults health"
- "bad enough to make appi"
- "body's control system + time"

Restores health

Healthcare Encounter
- "Black box"

"Health" or physiologic reserve

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

Observations (not recorded in EMR)

Observations about EMR use

Outcomes (not typically measured)

Collect more data: Order Tests

Ex: Laboratory, Radiology, Pathology/ exploratory procedures

Medications (orders, prescriptions)

Procedures (surgery, PT/OT, amputation)

Unregulated changes (diet)

Decisions to do something $d_1$

Clinical Decision Support & Information Synthesis Tools

EMR: CPOE, eRx

Actions (not recorded in EMR)

Observations about CDS effectiveness $\theta_2$

Interventions $i_1$

Outcomes (directly measured by health care delivery system)

"Health" or physiologic reserve

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

Collect more data: Order Tests

Outcomes (not typically measured)

Medications (orders, prescriptions)

Procedures (surgery, PT/OT, amputation)

Unregulated changes (diet)

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

Collect more data: Order Tests

Ex: Laboratory, Radiology, Pathology/ exploratory procedures

Medications (orders, prescriptions)

Procedures (surgery, PT/OT, amputation)

Unregulated changes (diet)