Advancing Clinical and Translational Research with Informatics at the University of Kansas Medical Center

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

This project is supported in part by NIH grant UL1TR000001.
• How we currently engage users
• Developing an Informatics Plan for the University of Kansas Medical Center
  – Looking for win-wins
  – What is needed for general research infrastructure versus informatics research
  – Repurposing CTSA efforts to Cancer Center NCI designation
• Realizing Return and Future Directions

Disclaimer: Clinical research biased
We have tools and expertise to manage data and convert it into information

**REDCap and CRIS** – 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 July 10.
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? Medical Informatics

**Current Goal:** Build Hatchery, Manage the Fishery
Second Goal: If you need help fishing, hire a guide from Medical Informatics

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

Photo Credit: S. Klashill
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

• 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?
• Prevent the wrong people (statistician/student) from entering/ changing data?

• Hospital/Clinic is making me use this Electronic Medical Record and I get nothing in return...

Little White Salmon River, Washington State, last Summer in July
Sometimes, You’re willing to enter data/buy fish:

**REDCap: Research Electronic Data Capture**

- [https://redcap.kumc.edu](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 and Surveys

- For consultation and to move project to production: Register your project with us so we can keep track of your request.
  - [http://frontiersresearch.org/frontiers/biomedical-informatics](http://frontiersresearch.org/frontiers/biomedical-informatics)

- Check out other institutions using REDCap and possibly borrow from the master library.
  - [http://www.project-redcap.org/](http://www.project-redcap.org/)
REDCap Case Report Form Example

University of Kansas Medical Center
University of Kansas Medical Center

KU-ADC

NACC - A1 Subject Demographics

Editing existing ADC Subject ID 1 (John Doe, RED_1)

ADC Subject ID

NOTE: This form is to be completed by intake interviewer per ADC scheduling records, subject interview, medical records, and proxy informant report (as needed). For additional clarification and examples, see UDS Coding Guidebook for Initial Visit Packet, Form A1. Check only one box per question.

Enrolled in NACC MDS

* must provide value

- Yes
- No

Primary Reason for coming to ADC

* must provide value

- Participate in research study
- Clinical evaluation
- Other
- Unknown

Principal Referral source

Clinician

Presumed disease Status at enrollment

Control/normal

Presumed Participation

Initial Evaluation only

ADC enrollment type

Subject’s month of birth

Subject’s year of birth

Subject’s sex

- Male
- Female

Does the subject report being Hispanic/Latino ethnicity (i.e., having origins from a mainly Spanish-speaking Latin American country), regardless of race?

- Yes
- No
- Unknown

What does subject report as his/her primary race

What additional race does subject report?

What additional race beyond what is indicated above in questions, does subject report?

Subject’s primary language

Informant’s years of education (report achieved level using the codes below; if an attempted level is not completed, enter the number of years attended). High school/GED = 12; Bachelors degree = 16; Master’s degree = 18; Doctorate = 20 years.
REDCap Survey: Think SurveyMonkey
Option Two: CRIS/Velos

REDCap Disclaimer

- For clinical trials, CRIS/Velos may be a better fit
  - Multiple years of experience
  - CRIS team builds for you with biostatistics review
  - Budget for CRIS team and biostatistics explicitly
- "Investigator driven" REDCap only works if you, the Principal Investigator, takes responsibility for your data
  - Scalability: informatics provides consultation and responsibility for technical integrity; *not your dictionary or data entry.*
    - Underwritten by CTSA, but you “feed and talk to your fish”
  - Middle model where informatics can build for you in REDCap.
    - Again, you budget for our team’s time
REDCap: think Fish Tank you manage

http://www.flickr.com/photos/wicked/185270913/lightbox/
CRIS: Vancouver Aquarium. We feed fancy fish; Biostatistics cooks tasty meal
I want to go fishing, not fill a fish tank (REDCap)  
Use HERON: a managed fishery

Bonneville Hatchery: Trout, Salmon, Sturgeon, Columbia River, Oregon
Aim #2: 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.

- **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
• Check http://informatics.kumc.edu/work/blog for latest status
The i2b2 “Fishing Rod”: build Diabetes cohort

Drag concepts from upper left into panels on the right

Types of “fish” in folders
i2b2 : **AND** in Frontiers Research Registry

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 Timelines
The dream: landing the big one

http://www.oregon.com/columbia_gorge_attractions/bonneville_hatchery
Without getting bit
Developing an Informatics Plan for KUMC

NIH Clinical Translational Science Awards are somewhat an anti-grant

- Provide a **portal for investigators to access** clinical and translational research **resources**, track usage and outcomes, and provide informatics consultative services.
- Create a **platform**, HERON (Healthcare Enterprise Repository for Ontological Narration), to **integrate clinical and biomedical data** for translational research.
- Advance medical innovation by linking **biological tissues** to **clinical phenotype** and the pharmacokinetic and pharmacodynamic **data generated by research cores** in phase I and II clinical trials (addressing T1 translational research).
- **Leverage** an active, engaged statewide **telemedicine** and Health Information Exchange (HIE) effort to enable community based translational research (addressing T2 translational research).
How do you know you’re winning?

- Added REDCap self service model to complement Velos Clinical Trial Management System
- For the last 2 quarters KUMC has some of the highest growth in new users nationally (> 700)
- Easiest thing you can do
- Low barrier to entry, no extra passwords low management cost (<1 FTE)
- But, people are free to screw up their research
Moving to FusionIO storage memory platform (August 2011) improved performance, but after upgrading to i2b2 1.6, query times increased significantly.

**Hardware/Software:** SUSE Linux Enterprise Server 11 (x86_64), Oracle 10g, 70G of RAM, 12 CPUs, Fusion-IO 1.28 TB IoDrive Duo storage tier
HERON: Current Contents

- 755 million facts
- 1.9 million patients but...
  - Most are just old administrative registrations
Richness of Phenotype is the Goal. Example: Frontiers Participant Registry

Frontiers Participant Data Richness Today

*All Frontiers Participants have Diagnosis and Procedure Data.
Please note that the full text is not visible in the image. However, the visible part of the text includes:

- **Dedicated Coordinator.** 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**
Supporting National Cancer Institute Cancer Center Designation

Incorporate Clinical, Administrative, Research Datasources

- Inpatient and outpatient electronic medical records (Epic)
- Professional Services Billing and Scheduling (GE IDX)
- KUCC Biospecimen Shared Resource Samples Database
- Hospital (KUH) Tumor Registry (NAACCR format)
- Social Security Death Master File (NIST format)
- Technical Charges from hospital and clinics (UHC validated format)
- Research Data Capture (REDCap)
- Clinical Research Information System (Velos)

HERON's current contents with new Cancer Center centric data in green

- Demographics (master patient index)
- Race/Ethnicity
- Laboratory Results
- Nursing observations/vital signs
- Clinical Diagnoses (ICD9)
- Medications (dispensed, ordered, home meds)
- Procedure charges (CPT)
- Outpatient Billing diagnoses (ICD9)
- Specimen collected
- Tumor Staging and Grade
- Diagnosis and Treatment
- Survival and Progression
- Death per Social Security Administration
- MSDRG
- Technical Charge Diagnoses ICD9
- Service line, AHRQ quality measures
- Triple Negative Breast Cancer Registry

initial pilot scheduled for release this month

Status as of July 15, 2012

http://informatics.kumc.edu/work/wiki/HeronProjectTimeline#March2012Planning
- contains current plan for next several monthly releases

FRONTIERS
THE HEARTLAND INSTITUTE FOR CLINICAL AND TRANSLATIONAL RESEARCH
Idealized HERON Cancer Center Research Workflow

1. See what we have
2. Define a cohort
3. Conduct Analysis
4. Resulting Plot

Cancer Survival: Obese (BMI > 30) Diabetic Breast

FRONTIERS
THE HEARTLAND INSTITUTE FOR
CLINICAL AND TRANSLATIONAL RESEARCH
Realizing Return and Future Directions

• Field of Dreams: if you build it, will they come?
  – Decreasing bioinformatics costs versus pricey EMR and extraction
  – **Engagement** to maximize return: Enterprise and Self-Service versus Fee for Service financial model

• Clinics/Hospital as Informatics “Lab” versus Direct Service
  – “Pathologist”: Diagnosing what’s happening with EMR and outcomes
  – “Surgeon”: apply informatics to influence clinical behavior and care
  – Service model for partnership with informatics versus standard consultant vendor model for hospitals? **Estimated win-win return**

• Expansion/Collaboration City, State, and Beyond?
  – Health care happens everywhere once you’re a decent catchment area
  – Unique assets may be less about money than people/political/cultural
  – Sharing to create environment for research? example: SHRINE project