Medical Informatics: A Fishing Story

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
September 20, 2012
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 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/
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/)
University of Kansas Medical Center

KU-ADC

NACC - A1 Subject Demographics

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

ADC Subject ID 1

**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**
  - Yes
  - No

- **Primary Reason for coming to ADC**
  - Participate in research study
  - Clinical evaluation
  - Other
  - Unknown

- **Specify**
  - [text]

- **Principal Referral source**
  - Clinician

- **Assumed 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/Latina 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 attempt level is not completed, enter the number of years attended). High school/GED = 12; Bachelor’s degree = 16; Master’s degree = 18; Doctorate = 20 years**
REDCap Survey: Think SurveyMonkey
REDCap Disclaimer

• “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/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
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.
• 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](http://informatics.kumc.edu/work/blog) for latest status
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
### Person #1774 f Non-hispanic
- **BMI (Calculated):** 30.070
- **Glucose (2011):** 110 (96% of goal) 100-110 (N=1385)
- **Hemoglobin A1c (2011):** 5.4 (N=23)

### Person #1809 f Non-hispanic
- **BMI (Calculated):** 30.070
- **Glucose (2010):** 110 (96% of goal) 100-110 (N=1385)
- **Hemoglobin A1c (2011):** 5.4 (N=23)

### Person #21253 f Non-hispanic
- **BMI (Calculated):** 30.070
- **Glucose (2010):** 110 (96% of goal) 100-110 (N=1385)
- **Hemoglobin A1c (2011):** 5.4 (N=23)

### Person #13473 f Non-hispanic
- **BMI (Calculated):** 30.070
- **Glucose (2010):** 110 (96% of goal) 100-110 (N=1385)
- **Hemoglobin A1c (2011):** 5.4 (N=23)

### Person #5138 m Non-hispanic
- **BMI (Calculated):** 30.070
- **Glucose (2010):** 110 (96% of goal) 100-110 (N=1385)
- **Hemoglobin A1c (2011):** 5.4 (N=23)

### Person #25806 f Non-hispanic
- **BMI (Calculated):** 30.070
- **Glucose (2010):** 110 (96% of goal) 100-110 (N=1385)
- **Hemoglobin A1c (2011):** 5.4 (N=23)
The dream: landing the big one

http://www.oregon.com/columbia_gorge_attractions/bonneville_hatchery

Catch the data for JAMA, NEJM publication
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).
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
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

• “Clinical Research” Informatics Next Steps
  – Decreasing bioinformatics costs versus pricey EMR and extraction
  – Engagement to maximize return: Enterprise self-service model with assistance by coordinator. Funding for complex or new data sources
  – Descriptive Statistics and Knowledge Discovery tools assist in data exploration and hypothesis generation
  – Build tools harmonize terminologies/clinical content
  – Tools for enhanced data delivery and preprocessing
  – Sharing to create environment for research example: SHRINE project

• Clinical Informatics: Service Clinics/Hospital, create ‘Lab’
  – “Pathologist”: Diagnosing what’s happening with EMR and outcomes
  – “Surgeon”: apply informatics to influence clinical behavior and care
    • My expertise from Vanderbilt