How Medical Informatics and HERON Can Help Your Research in the Department of Medicine?

Russ Waitman
Division of Medical Informatics
Department of Biostatistics
December 2, 2011
Outline for Today’s Presentation

- Informatics CTSA Aims: focus on storing and getting
  - Tools for storing information: REDCap
  - Tool for viewing/getting information: HERON/i2b2
    - Oversight Process
    - Milestones
- Upcoming Clinical Data in HERON
- Future Functionality and Integration
  - 1.6 release versus current 1.4: visits and modifiers
- Open time to explore hypotheses
National Clinical and Translational Science Award (CTSA) Objectives:

The purpose of this initiative is to assist institutions to forge a uniquely transformative, novel, and integrative academic home for Clinical and Translational Science that has the consolidated resources to:

1) captivate, advance, and nurture a cadre of well-trained multi- and inter-disciplinary investigators and research teams;

2) create an incubator for innovative research tools and information technologies; and

3) synergize multi-disciplinary and inter-disciplinary clinical and translational research and researchers to catalyze the application of new knowledge and techniques to clinical practice at the front lines of patient care.
Frontiers Biomedical Informatics Aims

1. Provide a portal for investigators to access clinical and translational research resources, track usage and outcomes, and provide informatics consultative services.

2. Create a platform, HERON (Healthcare Enterprise Repository for Ontological Narration), to integrate clinical and biomedical data for translational research.

3. 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).

4. Leverage an active, engaged statewide telemedicine and Health Information Exchange (HIE) effort to enable community based translational research (addressing T2 translational research).
REDCap: Research Electronic Data Capture

- [https://redcap.kumc.edu](https://redcap.kumc.edu)
  - It uses the same username and password as your KUMC email.
  - 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 make keep track of your request.
  - [http://biostatistics.kumc.edu/projectReg.aspx](http://biostatistics.kumc.edu/projectReg.aspx)
  - After you register your project, a CRIS team member, likely Kahlia Ford will get in touch with you.

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

- For clinical trials, CRIS/Velos may be a better fit
  - Multiple years of experience
  - CRIS team builds for you with biostats review
  - Budget for CRIS team and biostats explicitly

- “Investigator driven” REDCap works if PI takes responsibility for data
  - Scalability: informatics provides consultation and responsibility for technical integrity; *not your dictionary*.
    - Underwritten by CTSA right now
  - Or middle model where informatics can build for you in REDCap.
    - Again, you budget for our team’s time
## REDCap Case Report Form Example

### NACC - A1 Subject Demographics

**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.

<table>
<thead>
<tr>
<th>Field</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrolled In NACC MDS</td>
<td>Yes, No</td>
</tr>
<tr>
<td>Primary Reason for coming to ADC</td>
<td>Participate in research study, Clinical evaluation, Other, Unknown</td>
</tr>
<tr>
<td>Principal Referral source</td>
<td>Clinician, test</td>
</tr>
<tr>
<td>Presumed disease Status at enrollment</td>
<td>Control/normal</td>
</tr>
<tr>
<td>Presumed Participation</td>
<td>Initial Evaluation only</td>
</tr>
<tr>
<td>ADC enrollment type</td>
<td></td>
</tr>
<tr>
<td>Subject's month of birth</td>
<td></td>
</tr>
<tr>
<td>Subject's year of birth</td>
<td></td>
</tr>
<tr>
<td>Subject's sex</td>
<td>Male, Female</td>
</tr>
<tr>
<td>Does the subject report being Hispanic/Latino ethnicity (i.e. having origins from a mainly Spanish-speaking Latin American country), regardless of race?</td>
<td>Yes, No, Unknown</td>
</tr>
<tr>
<td>What does subject report as his/her primary race</td>
<td></td>
</tr>
<tr>
<td>What additional race does subject report?</td>
<td></td>
</tr>
<tr>
<td>What additional race beyond what is indicated above in questions, does subject report?</td>
<td></td>
</tr>
<tr>
<td>Subject's primary language</td>
<td></td>
</tr>
</tbody>
</table>

---

*University of Kansas Medical Center*

*KU-ADC*

*The REDCap Consortium | Data REDCap | University of Kansas Medical Center*
REDCap Survey Example
Aim #2: Create a data “fishing” platform:
HERON, https://heron.kumc.edu

- Develop business agreements, policies, data use agreements and oversight.

- Implement open source NIH funded (i.e. i2b2) initiatives for accessing data.

- Transform data into information using the NLM UMLS Metathesaurus as our vocabulary source.

- Link clinical data sources to enhance their research utility.
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
Big, well sorted, healthy trout
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/
The goal: a tasty publication

Photo Credit: Steve Velo
http://www.flickr.com/photos/juniorvelo/259888572/
Develop business agreements, policies, data use agreements and oversight.

- September 2010 the hospital, clinics and university signed a master data sharing agreement to create the repository.
  - Executive Committee – decides organization/systems expansion
  - Data Request Oversight Committee – guides implementation and approves/monitors use.

- Use Cases:
  - After signing a system access agreement, cohort identification queries and view-only access is allowed but logged and audited
  - Requests for de-identified patient data, while not deemed human subjects research, are reviewed.
  - Identified data requests require approval by the Institutional Review Board prior to data request review.
  - Contact information from the Frontiers Participant Registry have their study request and contact letters reviewed by the Participant and Clinical Interactions Resources Program.
Current Functionality

• Single sign-on using your email username
• Must be on campus or use VPN
• Real-time check for current human subjects training
• System Access Agreements, Data Use Agreements and Review Processes implemented in HERON with web pages for monitoring system use
• Check [http://informatics.kumc.edu](http://informatics.kumc.edu) for latest status

• Later Demonstration
  • i2b2 and HERON tools
  • Flowsheet data is a nursing and health professions gold mine (but like a mine, you have to dig).
Implement NIH funded (i.e. i2b2) initiatives for accessing data.
i2b2: Count Cohorts
i2b2: Patient Count in Lower Left
i2b2: Ask for Patient Sets
i2b2: Analyze Demographics Plugin
i2b2: Demographics Plugin Result

- Patient Set: Migrain-Diagnos@22:11:52 [7-9-2010] [kwilten] [PATIENTSET_303]
- Patient Count: 89
- Age in Years:
  - 10-20: 1
  - 20-30: 7
  - 30-40: 17
  - 40-50: 22
  - 50-60: 24
  - 60-70: 14
  - 70-80: 3
  - 80-90: 1
- Sex:
  - F: 76
  - M: 13
i2b2: View Timeline
i2b2: Timeline Results
HERON De-identification Decisions

- **HIPAA Safe Harbor De-identification**
  - Remove 18 identifiers and date shifting by 365 days back
  - Resulting in non-human subjects research data but treated as a limited data set from a system access perspective.
    System users and data recipients agree to treat as a limited data set (acknowledging re-identification risk)

- **To be addressed:**
  - For now, we won’t add free text such as progress notes with text scrubbers (DeID, MITRE Identification Scrubber toolkit)
Data Sources for FY2012: Focus on Supporting Cancer Center Initiative

- HERON Executive Committee approval June 2011 for incorporating:
  - University Biospecimen Repository (Aim 3, Cancer Center)
  - Hospital Tumor registry (Aim 3, Cancer Center)
  - University REDCap and Velos Registries and Clinical Trials systems (Aim 3, Cancer Center)
  - Hospital billing ICD9, MS-DRG, Insurance Status
  - Social Security Death Master File (Aim 4, Cancer Center)
  - Cerner CoPath pathology system (Aim 3, Cancer Center)

- Also continue to extract and refine data from Epic EMR
Developing a Rich Description of our Population: Existing and Planned Data Sources for HERON. Existing sources shown in solid lines and planned in dashed, new sources this Fall/Winter in green.
An i2b2 query against HERON for currently supported cancer centric data sources

Any neoplasm ICD9 diagnosis (106,000 patients) and a WBC count (121,000) -> 44,000 distinct patients,
*require height (123,000) and weight (154,000) -> 35,000 patients,
•require Wong-Baker pain scale (84,000) -> 14468 patients,
•Body Temperature (158,000) -> 14463 patients,
•Surgical Pathology Procedures CPT (85,000) -> 12446 patients,
Finally selective serotonin 5-HT3 antagonist antiemetics -> 8517 patients
With our improved hardware (Fusionio memory cards), the cohort size is returned in 15 seconds for this 8 group query.
Biospecimen Shared Resource Integration
Adding Social Security Death Master File

- Have Death status on approximately 90 million people.
  - Contains Social Security Number, Name, Date of Birth, Date of Death, Place of Death
  - Monthly update file from ntis; will sync with releases

- **Released Friday, September 23, 2011**
  - Matching on SSN plus DOB
  - 177,706 of our 1.8 million people noted as deceased according to Social Security Administration versus 23,850 from hospital systems.
New Functionality: IRB and i2b2 1.6

- Moving beyond counting to line item data review
  - In August, Karen Blackwell Privacy officials agreed to allow timeline access under current system access agreement
  - Released Friday, September 23, 2011

- i2b2 version 1.5: DataMart Request Form to facilitate our Data Use Agreement

- i2b2 version 1.6: Visit enabled queries

- i2b2 Modifiers with i2b2 version 1.6
  - Will have to redo ELT to take advantage
  - Targeting Christmas 2011 and diagnoses as first example
KUH Tumor Registry

- Validated Outcomes and Observations
  - Tumors, Nodes, Metastasis (TNM) on complete cases
  - Untapped investment: 7 cancer registrars (Tim Metcalf)
  - ~60,000 cases, data since 1950s

- North American Association of Central Cancer Registries (NAACCR) file format
  - Will build on work at other NCI designated i2b2 users
    (Group Health Cooperative in Seattle, Kimmel Cancer Center in Philadelphia have shared their code/metadata with us)
  - John Keighley providing invaluable expertise

- Later, supplement with additional treatment information not in NAACCR file

Draft released Tuesday, November 22, 2011
What do Visits and Modifiers Offer?

- **Visits:**
  - I want to know the patient had the lab and the medication in the same episode of care.
  - Conceptually, i2b2 has had a table for the visit dimension but the software never exploited the data.

- **Modifiers:**
  - Is it a billing diagnosis or from the problem list? Is it a primary or secondary?
  - How do I represent all parts of a medication order (dose, route, frequency)?
Constrain observations to the same visit

Murphy SN et al, https://www.i2b2.org/events/slides/i2b2_OpeningTalk_20110628_Murphy.pdf
i2b2 Modifiers in the User Interface

Murphy SN et al, https://www.i2b2.org/events/slides/i2b2_OpeningTalk_20110628_Murphy.pdf
i2b2 Modifiers in the User Interface

Murphy SN et al, https://www.i2b2.org/events/slides/i2b2_OpeningTalk_20110628_Murphy.pdf
i2b2 Modifiers in the User Interface

Murphy SN et al, https://www.i2b2.org/events/slides/i2b2_OpeningTalk_20110628_Murphy.pdf
i2b2 Modifiers in the User Interface

Murphy SN et al, https://www.i2b2.org/events/slides/i2b2_OpeningTalk_20110628_Murphy.pdf