Molding and Maximizing the LabKey Platform for Clinical Translational Research

Anthony Corbett
University of Rochester Medical Center
2019 LabKey User Conference
• LabKey at URMC

• Philosophy

• Strategy
  • Clinical Data Capture
  • Specimen Management
  • Experimental Data
  • Reporting and Visualization

• Future Work
Multi-Center grants

Industry funded research projects

Biobanking

Departmental/Division wide

Reagent databases
Research Areas

Infectious diseases

Autoimmune diseases

Lung development

Cancer

Clinical Interventions in Autism

Public Health (smoking/vaping)

Maternal and Children’s Health
Overview

Bio-Lab-Informatics-System (BLIS)
- Over 90 observational/registry studies, since 2010
- ~16,000 participants
- ~450,000 specimens (40 different primary specimen types)
- ‘omics and assay data integration

Biorepository for Investigation of Neonatal Diseases of the Lung (BRINDL)
- ~72,000 samples
- ~200 donors
- Imaging and ‘omics integration
• Improve Research Team Collaboration & Productivity
  • Participate as member of research teams
  • User, processes and workflow focused

• Enable “Real-time” Data
  • Visualizations, Analytics, Exploration, Sharing

• Enhanced Data Stewardship & Security (FISMA)

• Innovate Incorporating Best Tools and Practices
“Organizations which design systems ... are constrained to produce designs which are copies of the communication structures of these organizations.”

- Conway’s Law

Thus, the structure of a system will reflect the social boundaries of the organization(s) that produced it, across which communication is more difficult.
Factors driving LabKey adoption at URMC

Single Lab -> Collaborative Science

Single Purpose -> Discovery and Reuse
Clinical Data
1. Using the **redcap module** to import clinical data in CDISC ODM format

2. Use the **data integration module** to write SQL and ETL processes
   - Restructure and subset the data
   - Calculated fields
   - QC metrics
Custom REDCap Interactions

URMC developed module

Redcap API Service
- Import and Export Records
- Download files
- Run RedCAP Reports
- Download Events, Arms, Event-Form Mappings.

Custom ETL task definitions
- Export Records Task
- Download Files into DataClass and process in LabKey
URMC developed EDC framework built on LabKey JavaScript API
Specimen Inventory
Excel spreadsheets

LDMS

BSI
• Specimen accession and processing
• Freezer and location management
• Integration with BarTender
  • Barcoding
  • label printing
One way: BSI -> LabKey
Pulled Nightly with
Custom Availability Rules

Want two way: LabKey <-> BSI
Specimen Requests -> Requisitions
Instrument and Assay Data
### RNA QC Summary

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>fC140_1</td>
<td></td>
<td>26.8</td>
<td>32.36</td>
<td>867.24</td>
<td>22.86</td>
<td>612.79</td>
<td>2.12</td>
<td>1.4</td>
<td>9.7</td>
</tr>
<tr>
<td>fC140_2</td>
<td></td>
<td>26.8</td>
<td>5.55</td>
<td>148.74</td>
<td>1.57</td>
<td>42.21</td>
<td>1.67</td>
<td>0.18</td>
<td>7.6</td>
</tr>
<tr>
<td>fC140_3</td>
<td></td>
<td>26.8</td>
<td>9.57</td>
<td>256.47</td>
<td>5.51</td>
<td>147.77</td>
<td>2.25</td>
<td>0.87</td>
<td>9.7</td>
</tr>
<tr>
<td>FP_A_Plus_1</td>
<td></td>
<td>26.8</td>
<td>13.29</td>
<td>356.17</td>
<td>3.55</td>
<td>95.17</td>
<td>2.26</td>
<td>0.03</td>
<td>9.9</td>
</tr>
<tr>
<td>FP_A_Plus_2</td>
<td></td>
<td>26.8</td>
<td>3.3</td>
<td>88.44</td>
<td>1.03</td>
<td>27.76</td>
<td>2.0</td>
<td>0.45</td>
<td>8.4</td>
</tr>
<tr>
<td>FP_A_Plus_3</td>
<td></td>
<td>26.8</td>
<td>3.38</td>
<td>90.58</td>
<td>1.32</td>
<td>35.62</td>
<td>2.76</td>
<td>0.43</td>
<td>9.4</td>
</tr>
</tbody>
</table>

**BioAnalyzer and Nanodrop QC**
## Illumina Sequencing Run and Results

<table>
<thead>
<tr>
<th>Specimen ID</th>
<th>Primary Type</th>
<th>Derivative Type</th>
<th>Treatment</th>
<th>Flow Cell Id</th>
<th>Lane</th>
<th>Index</th>
<th>Sample Ref</th>
<th>Read 1 Sequence File</th>
<th>Assay Id</th>
<th>Run Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>A160TRJ7-12</td>
<td>BLD</td>
<td>XPA</td>
<td>CD4</td>
<td>C7R1ACXX</td>
<td>1</td>
<td>CGAGGCTG-CTCTCTAT</td>
<td>Human</td>
<td>sortedBlood_RNASEq/20150605_CD4/seqs/raw/10_CD4_R1.fastq.gz</td>
<td>CD4 Season 3</td>
<td>Season 3</td>
</tr>
<tr>
<td>B160TSH5-11</td>
<td>BLD</td>
<td>XPA</td>
<td>CD4</td>
<td>C7R1ACXX</td>
<td>1</td>
<td>AAGGGGCA-TATCTCTCT</td>
<td>Human</td>
<td>sortedBlood_RNASEq/20150605_CD4/seqs/raw/11_CD4_R1.fastq.gz</td>
<td>CD4 Season 3</td>
<td>Season 3</td>
</tr>
<tr>
<td>H160TSJD-12</td>
<td>BLD</td>
<td>XPA</td>
<td>CD4</td>
<td>C7R1ACXX</td>
<td>1</td>
<td>GTAGAGGA-AGATAGA</td>
<td>Human</td>
<td>sortedBlood_RNASEq/20150605_CD4/seqs/raw/12_CD4_R1.fastq.gz</td>
<td>CD4 Season 3</td>
<td>Season 3</td>
</tr>
<tr>
<td>J160TSZM-12</td>
<td>BLD</td>
<td>XPA</td>
<td>CD4</td>
<td>C7R1ACXX</td>
<td>1</td>
<td>TAAGGCGA-GTAAGGAG</td>
<td>Human</td>
<td>sortedBlood_RNASEq/20150605_CD4/seqs/raw/13_CD4_R1.fastq.gz</td>
<td>CD4 Season 3</td>
<td>Season 3</td>
</tr>
<tr>
<td>A160TT9Y-11</td>
<td>BLD</td>
<td>XPA</td>
<td>CD4</td>
<td>C7R1ACXX</td>
<td>1</td>
<td>CGTACTAG-ACTGCTCTA</td>
<td>Human</td>
<td>sortedBlood_RNASEq/20150605_CD4/seqs/raw/14_CD4_R1.fastq.gz</td>
<td>CD4 Season 3</td>
<td>Season 3</td>
</tr>
<tr>
<td>G160TT99-11</td>
<td>BLD</td>
<td>XPA</td>
<td>CD4</td>
<td>C7R1ACXX</td>
<td>1</td>
<td>AGGCAGAA-AAGGAGTA</td>
<td>Human</td>
<td>sortedBlood_RNASEq/20150605_CD4/seqs/raw/15_CD4_R1.fastq.gz</td>
<td>CD4 Season 3</td>
<td>Season 3</td>
</tr>
<tr>
<td>E160TQR-12</td>
<td>BLD</td>
<td>XPA</td>
<td>CD4</td>
<td>C7R1ACXX</td>
<td>1</td>
<td>TCCTGAGG-CTAACGCT</td>
<td>Human</td>
<td>sortedBlood_RNASEq/20150605_CD4/seqs/raw/16_CD4_R1.fastq.gz</td>
<td>CD4 Season 3</td>
<td>Season 3</td>
</tr>
<tr>
<td>F160TTPJ-11</td>
<td>BLD</td>
<td>XPA</td>
<td>CD4</td>
<td>C7R1ACXX</td>
<td>1</td>
<td>GGACTCCT-GCGTAAGA</td>
<td>Human</td>
<td>sortedBlood_RNASEq/20150605_CD4/seqs/raw/17_CD4_R1.fastq.gz</td>
<td>CD4 Season 3</td>
<td>Season 3</td>
</tr>
<tr>
<td>J160TV0G-11</td>
<td>BLD</td>
<td>XPA</td>
<td>CD4</td>
<td>C7R1ACXX</td>
<td>2</td>
<td>TAGGCATG-CTCTCTAT</td>
<td>Human</td>
<td>sortedBlood_RNASEq/20150605_CD4/seqs/raw/18_CD4_R1.fastq.gz</td>
<td>CD4 Season 3</td>
<td>Season 3</td>
</tr>
</tbody>
</table>
File share for QC and Analysis Reports
Dashboard and Reports
### Interactive Reports

#### Quality Control

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Visit Name</th>
<th>Functional Assessment Has Data</th>
<th>Lawton-Brody Has Data</th>
<th>Barthel Has Data</th>
<th>McH Has Data</th>
<th>Gf Has Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 Month followup</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2 Month followup</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>4 Month followup</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>6 Month followup</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>4 Month followup</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>6 Month followup</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>4 Month followup</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2 Month followup</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>4 Month followup</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>6 Month followup</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>4 Month followup</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

### Missing Records Plot

#### Subjects with Missing SLE ACR Criteria on Enrollment

- SLE
- SLE
- SLE
- SLE
- SLE
- SLE
- SLE
- SLE
- SLE
- SLE
Clinical Adjudication Summary
Integration with additional clinical data sources:
EMR and/or Clinical Data Warehouse

Better Integration with OMERO:
Image Viewing and Analysis

Managing Neuroimaging (fMRI) data:
Brain Imaging Data Structure (BIDS)

Integration with computational workflow engines:
Galaxy and CWL/WDL
Acknowledgements

**RDIA Team**
Jeanne Holden-Wiltse, MPH, MBA, Director
Anthony Corbett, MS, Technical Lead
Sanjukta Bandyopadhyay, MS
Jennifer Dutra
Jill Payne
Alicia Tyrell, MBA
Andrew Watts, MA
Jeff Williams
Zidian Xie, PhD
Mike Conklin
Kim Kaukeinen

**LabKey Corporation**
**CTSI Informatics**
**Center for Integrated Research Computing**

**Major Funding**