The National Consortium on Alcohol and Neurodevelopment in Adolescence (NCANDA): A Framework Supporting Neuroimaging Data Integration and Analysis

Nolan Nichols, Ph.D.\textsuperscript{1,2}
Weiwei Chu, M.S.\textsuperscript{1}
Kilian Pohl, Ph.D.\textsuperscript{1,2}

\textsuperscript{1}Center for Health Sciences, SRI International
\textsuperscript{2}Department of Psychiatry & Behavioral Sciences, Stanford
The Challenge

- Multi-Site
- Multi-Modal
- Longitudinal

Traditional Research Lab

National Consortium

- Five sites across the US
- Total of 808 enrolled participants
- Baseline, 1 year, and 2 year follow-ups
- Data collection initiated in 6 months
- Limited resources for development
- What are the system requirements?
**Requirements**

### Clinical and Neuropsychological Instruments
- Accommodate heterogeneous instruments
- Validate data capture protocols
- Maintain ongoing data quality
- Ensure longitudinal visit time windows
- Automate as much as possible
- What Neuroinformatics resources to reuse, circa 2012?

### Multimodal Imaging and Phantoms
- T1
- DTI
- rsfMRI
- ADNI/fBIRN
Overview and Approach

Sites collect:
- Demographic Information
- Clinical Data
- Neuropsychological Test Scores
- MRI
  - Anatomical
  - Diffusion
  - Functional
Clinical and Neuropsych Assessments

Extract → Transform → Load

REDCap Harmonization

Cross-sectional REDCap Project

Visit Classification & Scoring

Import to Central Repository

Longitudinal REDCap Project

<table>
<thead>
<tr>
<th>Subject ID</th>
<th>Visit Date and Notes  Baseline visit (Arm 1: Standard Protocol)</th>
<th>Visit Date and Notes Baseline visit (Arm 1: Standard Protocol)</th>
<th>Visit Date and Notes Baseline visit (Arm 1: Standard Protocol)</th>
<th>Visit Date and Notes 6-month follow-up (Arm 1: Standard Protocol)</th>
<th>Visit Date and Notes 1 y visit (Arm 1: Standard Protocol)</th>
<th>Visit Date and Notes 1 y visit (Arm 1: Standard Protocol)</th>
<th>Visit Date and Notes Baseline visit (Arm 1: Standard Protocol)</th>
<th>Visit Date and Notes 1 y visit (Arm 1: Standard Protocol)</th>
<th>Visit Date and Notes Baseline visit (Arm 1: Standard Protocol)</th>
<th>Visit Date and Notes 1 y visit (Arm 1: Standard Protocol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Record ID | Visit Information | Delayed Discounting, $1000 | Delayed Discounting, $100 |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2013-03-13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-02-07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-02-26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-02-18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-02-21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-02-22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-04-01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-03-15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-03-18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-07-02</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Multi-Modal Imaging

Sites Scans
T1, T2, DTI, rsfMRI, ADNI/fBIRN Phantoms

<table>
<thead>
<tr>
<th>Scans</th>
<th>Type</th>
<th>Series Desc</th>
<th>Usability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 T1 Localization Scan</td>
<td>ncanda-localizer-v1</td>
<td>ncanda-localizer-v1</td>
<td>unknown</td>
</tr>
<tr>
<td>2 T2 Calibration Scan</td>
<td>ncanda-calibration-v1</td>
<td>ncanda-calibration-v1</td>
<td>unknown</td>
</tr>
<tr>
<td>3 DTI Scan</td>
<td>ncanda-t1spgr-v1</td>
<td>ncanda-t1spgr-v1</td>
<td>usable</td>
</tr>
<tr>
<td>4 DTI Scan</td>
<td>ncanda-t2fse-v1</td>
<td>ncanda-t2fse-v1</td>
<td>usable</td>
</tr>
<tr>
<td>5 DTI Scan on Phantom</td>
<td>ncanda-dti6b500pepolar-v1</td>
<td>ncanda-dti6b500pepolar-v1</td>
<td>usable</td>
</tr>
<tr>
<td>6 DTI Scan on Phantom</td>
<td>ncanda-dti6b1000-v1</td>
<td>ncanda-dti6b1000-v1</td>
<td>usable</td>
</tr>
<tr>
<td>7 DTI Scan on Phantom</td>
<td>ncanda-grefieldmap-v1</td>
<td>ncanda-grefieldmap-v1</td>
<td>usable</td>
</tr>
<tr>
<td>8 DTI Scan on Phantom</td>
<td>ncanda-rsfmri-v1</td>
<td>ncanda-rsfmri-v1</td>
<td>usable</td>
</tr>
</tbody>
</table>

- Semi-automatic QA
- Neuroradiologist readings
- Hourly/nightly QA reports
- Event-based workflow to populate image processing pipeline
Data Analysis

Import to Pipeline

Light Weight Data Pipeline
http://www.nitrc.org/projects/lwdp

Anatomical
FreeSurfer Parcellation and Segmentation

Diffusion
FSL Tract-Based Spatial Statistics

Resting State
Nipype-based Preprocessing
Next Steps and Data Sharing

- Curated with REDCap data dictionaries
- Use Case for NIDM and BIDS Standards
- Adopt the NIDM process to curate data
- Develop object models of clinical, neuropsych, and imaging measures
- Demonstrate usage of NIDM for distributing NCANDA datasets

<table>
<thead>
<tr>
<th>Data Collection Sites</th>
<th>NCANDA Neuroinformatics Data Analysis Platform</th>
<th>Dataset Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site Form Collection</td>
<td>REDCap</td>
<td>NITRC</td>
</tr>
<tr>
<td>Site Imaging Collection</td>
<td>ClinicalNP</td>
<td>RDoC</td>
</tr>
<tr>
<td></td>
<td>Data Analysis &amp; Harmonization</td>
<td>OpenFMRI</td>
</tr>
<tr>
<td></td>
<td>Imaging Pipeline</td>
<td>Amazon webservice</td>
</tr>
</tbody>
</table>

**Data Dictionary to Ontology Mapping**

- Data Release
- NIDM Mapping
- Infer Task Concepts
- Query for Study Types
- Retrieve Documentation
- Specification Generation

**Example**

```
SELECT DISTINCT *
WHERE {
  ?project a nidm:Project .
  ?project dct:hasPart ?study .
  ?study dct:hasPart ?acquisition .
  ?study a ?study_type .
}
```
Conclusions and Lessons Learned

• The ecosystem of Neuroinformatics software for imaging studies is mature
• Neuroinformatics tools can be reused and extended to develop scalable Neuroinformatics platforms
• Early involvement of informaticians may be able to simplify system architecture using common platforms
• Version Control Systems are an innovative way to capture data asynchronously before a data management system implementation
• Demonstrated the reuse of neuroinformatics tools to provide data integration and analysis hub for a multi-site, longitudinal study on adolescent development

https://www.nitrc.org/projects/ncanda-datacore
Acknowledgements

National Institute on Alcohol Abuse and Alcoholism (NIAAA)
Data Analysis Component 5U01AA021697

Data Analysis Component
Weiwei Chu    Kilian Pohl
Kevin Cummins  *Torsten Rohlfing
Dongjin Kwon    Yong Zhang
Dolf Pfefferbaum

Data Sharing Task Force
Satra Ghosh
Chris Gorgolewski
David Keator
JB Poline
Jessica Turner
... and many others!
References


3. WebCNP: https://webcnp.med.upenn.edu/
5. Blaise: http://www.blaise.com
6. ePrime: http://www.pstnet.com/eprime.cfm

8. Subversion: https://subversion.apache.org/