Curating, Releasing, and Access NCANDA Data

Kilian M. Pohl
Adolf Pfefferbaum
Ehsan Adeli
Dongjin Kwon
Michael Hasak
Qingyu Zhao
Sara Benito
Weiwei Chu
SRI International

Edith V. Sullivan
Stanford University

kilian.pohl@sri.com  ♦  https://www.stanford.edu/~kpohl
Financial Interest Disclosure

Salary and Research Support of Kilian M Pohl, Ph.D.
• NCANDA collection sites upload data acquired in 831 study participants to the SIBIS portal
• SIBIS combines, curates, analyzes, and distributes the data using open source software.
Sites collect:

- Non Imaging Data
  - Demographic Information
  - Clinical Data
  - Neuropsychological Test Scores

- MRI
  - Structural
  - Diffusion
  - Functional
# Data Analysis

## Collection
- Site Form Collection
- Site Imaging Collection

## Import
- Convert & Upload
- Imaging Upload

## Analysis
- QC Upload
- Extract Scores
- Harmonize Data
- Group Analysis

### Tools
- LimeSurvey
- PenaCNP
- REDCap
- XNAT

### Data Sources
- Clinical/NP
- Neuroimaging
Semi-Automatic Quality Control

Automatically check dimension, ....
Visually check for image artifacts
Chiari 1 malformation

- bilateral tonsillar herniation with medullary distortion

right parietal cortical mass

- 95/833 adolescents = 11.4%
- 2 excluded from the study
Issues Reported To Github

Issues are resolved with collection sites

Kilian M. Pohl
Process Data

Import to Pipeline

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

Resting State
Diffusion
Anatomical

Nipype-based Preprocessing
Kilian M. Pohl

FSL Tract-Based Spatial Statistics

Volume Scores
Create Data Releases

• Contains text files with de-identified demographic and neuropsychological test scores, raw and derived imaging data, and composite scores
• Data dictionary describing variables
• Software used to generate the data
Create machine learning technology to identify brain regions impacted by regular alcohol consumption during adolescence
Share Data and Software

Upload results to the data repository Synapse
Store software via
- NITRC: https://www.nitrc.org/projects/ncanda-datacore
- Github: https://github.com/sibis-platform
Gaining Access to Data

Apply to NIAAA:

• A **cover letter** on the letterhead of the sponsoring institution at which the research study will be conducted.

• **Curriculum vitae** of the principal investigator and all co-investigators

• A 1-2 page **description of the proposed research**

• A completed **Data Distribution Agreement** with signatures of the principal investigator and an authorized representative of the sponsoring institution.

Gaining Access to Data

After approval from NIAAA

- SRI will request from PI their Synapse user name
- PI will be able to access data via https://www.synapse.org/ncanda

Synapse allows users to query and download data releases, which are linked to publications.
Data Sample

Data Dictionary

Score
Thank You

Dongjin Kwon, PhD
Ehsan Adeli, PhD
Qingyu Zhao, PhD
Michael Hasak
Sara Benito
Weiwei Chu