The NDAR Model for Publishing Findings in the Life Sciences

Dan Hall

Manager, National Database for Autism Research, NIMH

National Institute of Mental Health

National Database for Autism Research

- Joint initiative supported by NIMH, NICHD, NINDS, and NIEHS
  - Federal data repository
  - Contains data from human subjects related to autism (and control subjects)
  - Data are available to the research community through a not too difficult application process
  - Summary data are available to everyone with a browser at http://ndar.nih.gov

- Begun in late 2006, first data was received in 2008, significant data became available in 2012.

- Autism Interagency Coordinating Committee Recommendation called for 90% of all human subjects data to be shared to are harmonized.

- Currently has broad data available from over 77,000 subjects from demographic data, -omic (~400TB), clinical assessments, imaging data, eye tracking, and exposure.

NIH | National Institute of Mental Health

NDAR: Implementation

- NDAR data is harmonized:

  - Global Unique Identifier (GUID)

  - Data dictionaries (400+ measures across 60,000 discrete data elements)

  - Experiment tool for -omic, EEG, eye tracking, and fMRI

  - Study definition and results…

- NDAR has deep federation with important repositories controlled by others (Autism Speaks, the Simons Foundation, and Kennedy Krieger).

- NIH investigators are expected to share their data ongoing **"about"** research subjects (e.g. raw genomics, structural MRI, diagnostic measures) and **"findings"** at the time of publication.

- Over 100 studies have provided data, and more than 150 are now expected to share data closing in on the 90% community objective.

- While data sharing is successful, associating data directly to findings is now essential.

Results Reporting (NDAR Study)

Results Reported by:

- **Cohorts** - Grouping of data by subject by phenotype or pipeline

- **Measures** – List and share data by primary/secondary measure

- **Study Design** - Study description, controlled, observational, arms/comparisons, interventions, etc.

- **Analysis** - Keywords describing data analysis, i.e., data transformation, statistical methods, equipment, software, algorithms, etc.

- **Results**

The bar must be raised to clearly define and publish findings with the means for funders and publishers to easily validate that data supporting important findings are properly shared. Perhaps the NDS can help?

National Institute of Mental Health