**Autism Brain Imaging Data Exchange (ABIDE) Meta-Data Description**

**Goals**

1. Data Preprocessing, cleaning, missing patterns, etc.
2. EDA
3. QC assessment (use SOCR Chronbach’s alpha instrument validator, <http://socr.umich.edu/HTML5/SOCRAT/>, test first with Turkijy’s Eval data provided in app) compare consistency of different QC raters
4. Split data features into:
   1. Raw Data (quantitative vs. categorical) – clinical demographic, phenotypic
   2. QC data
   3. Derived data (neuroimaging local measures, majority of features) quantitative vs. categorical
   4. Analytics: exploratory analytics and ML classifiers (n=1098, k=2100), we will do some advanced model-free predictions and forecasting of clinically relevant measures (e.g., Dx\_Category, Dx, Site, DSM, IQ, etc.) Need to employ knockoff filtering to reduce FDR error rate of selected features (<http://www.socr.umich.edu/people/dinov/2017/Spring/DSPA_HS650/notes/17_RegularizedLinModel_KnockoffFilter.html>)
   5. Other methods

**Neuroimaging Data Processing**

Quality assessment (QA) columns represent automatically derived metrics prefixed by ***anat\_***, ***func\_***, and manual quality assessment prefixed by ***qc\_***.

**Automated QA Measures**

These columns reflect automated metrics where outliers may be identified by a statistical procedure (e.g., 2\*\sigma) and include:

**Anatomical measures**

**Contrast to Noise Ratio [anat\_cnr]**: mean of the gray matter values minus the mean of the white matter values, divided by the standard deviation of the air values 1.

**Entropy Focus Criterion [anat\_efc]:** Shannon’s entropy is used to summarize the principal directions distribution, higher energy indicating the distribution is more uniform (i.e., less noisy) 2.

**Foreground to Background Energy Ratio [anat\_fber]**: Mean energy of image values (i.e., mean of squares) within the head relative to outside the head.

**Smoothness of Voxels [anat\_fwhm]**: The full-width half maximum (FWHM) of the spatial distribution of the image intensity values in terms of voxels (e.g., a value of 3 implies smoothness of 3 voxels) 3.

**Percent of Artifact Voxels [anat\_qi1]:** The proportion of voxels with intensity corrupted by artifacts normalized by the number of voxels in the background 4.

**Signal to Noise Ratio [anat\_snr]**: The mean of image values within gray matter divided by the standard deviation of the image values within air (i.e., outside the head) 1.

**Functional measures**

**Entropy Focus Criterion [func\_efc]:** Shannon’s entropy is used to summarize the principal directions distribution, higher energy indicating the distribution is more uniform (i.e., less noisy) 2. Uses mean functional.

**Foreground to Background Energy Ratio [func\_fber]:** Mean energy of image values (i.e., mean of squares) within the head relative to outside the head. Uses mean functional.

**Smoothness of Voxels [func\_fwhm]:** The full-width half maximum (FWHM) of the spatial distribution of the image intensity values. Uses mean functional.

**Standardized DVARS [func\_dvars]:** The spatial standard deviation of the temporal derivative of the data, normalized by the temporal standard deviation and temporal autocorrelation 56. Uses functional time-series.

**Fraction of Outlier Voxels [func\_outlier]:** The mean fraction of outliers found in each volume using 3dTout command in AFNI (http://afni.nimh.nih.gov/afni) 7. Uses functional time-series.

**Mean Distance to Median Volume [func\_quality]:** The mean distance (1 – spearman’s rho) between each time-point’s volume and the median volume using AFNI’s 3dTqual command (<http://afni.nimh.nih.gov/afni>)7. Uses functional time-series.

**Mean Framewise Displacement (FD) [func\_mean\_fd]:** A measure of subject head motion, which compares the motion between the current and previous volumes. This is calculated by summing the absolute value of displacement changes in the x, y and z directions and rotational changes about those three axes. The rotational changes are given distance values based on the changes across the surface of a 50mm radius sphere 58. Uses functional time-series.

**Number FD greater than 0.2mm [func\_num\_fd]**: The number of frames or volumes with displacement greater than 0.2mm. This is not shown below. Uses functional time-series.

**Percent FD greater than 0.2mm [func\_perc\_fd]:** The percent of frames or volumes with displacement greater than 0.2mm. Uses functional time-series.

**Ghost to Signal Ratio [func\_gsr]:** A measure of the mean signal in the ‘ghost’ image (signal present outside the brain due to acquisition in the phase encoding direction) relative to mean signal within the brain. Uses mean functional.

**Manual QA measures**

Manual inspection of the data was carried out by three independent raters. Rater 1 examined the general quality of the data focusing on the preprocessed functional data and derivatives. Raters 2-3 examined the quality of the raw anatomical and functional data. Ratings are given as OK and fail for raters 1 and 3, and OK, maybe, and fail for rater 2. Each rater also may have taken notes related to their ratings, thus there is a notes column associated with each rating column.

**References**

Magnotta, V. A., & Friedman, L. (2006). Measurement of signal-to-noise and contrast-to-noise in the fBIRN multicenter imaging study. Journal of Digital Imaging, 19(2), 140-147. ↩ ↩2

Farzinfar, M., Dietrich, C., Smith, R. G., Li, Y., Gupta, A., Liu, Z., & Styner, M. A. (2012, May). Entropy based DTI quality control via regional orientation distribution. In Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on (pp. 22-25). IEEE. ↩ ↩2

Friedman, L., Stern, H., Brown, G. G., Mathalon, D. H., Turner, J., Glover, G. H., … & Potkin, S. G. (2008). Test–retest and between‐site reliability in a multicenter fMRI study. Human brain mapping, 29(8), 958-972. ↩

Mortamet, B., Bernstein, M. A., Jack, C. R., Gunter, J. L., Ward, C., Britson, P. J., … & Krueger, G. (2009). Automatic quality assessment in structural brain magnetic resonance imaging. Magnetic Resonance in Medicine, 62(2), 365-372. ↩

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. & Petersen, S. E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59, 2142-2154 (2012). ↩ ↩2

Nichols, T. (2012, Oct 28). Standardizing DVARS. Retrieved from http://blogs.warwick.ac.uk/nichols/entry/standardizing\_dvars.

Cox, R.W.. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. Computers and Biomedical Research, 29:162-173, 1996.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage, 17(2), 825-841.

**ABIDE Metrics** <http://preprocessed-connectomes-project.org/quality-assessment-protocol/#normative-metrics>

**FSL morphometrics**:

* <https://surfer.nmr.mgh.harvard.edu/fswiki/MorphometryStats>
* <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2901819/>

**LPBA40-BrainParser Nomenclature:**

* <http://wiki.socr.umich.edu/index.php/SOCR_Data_N46_TBI_ROI_Volumes>
* <http://www.loni.usc.edu/atlases/Atlas_Methods.php?atlas_id=20>

**Common abbreviations**

BA = Brodmann area <https://en.wikipedia.org/wiki/Brodmann_area>

Thick = cortical thickness

LH – left hemkisphere

RH = right hemisphere

Curv = curvature

Surf = surface

Ind = index

Gaus = Gaussian

Std = standardized/standard