GENUS Consortium Meeting - International Congress On Schizophrenia Research
Orlando, FL – April 22, 2013 – 5:30 to 6:30 pm

Overview of project aims:
Investigate schizophrenia risk genes for involvement in brain dysfunction as indexed by neuropsychological deficits and neuroanatomical (structural MRI, DTI) abnormalities to understand the role of risk genes in disease.

- Focus on genes with genome-wide significant association with scz from prior GWAS. Now ~100 genes from the PGC2 scz mega-analysis. Consider analyzing as pathways or choose subset to reduce number of tests?
- Also top genes from GWAS of cognition (e.g., COGENT), brain structure (e.g., ENIGMA), etc.
- Investigate polygenic factor (from PGC2) in variation in neuropsych and imaging measures.
- Multivariate analysis of phenotypes, within or across neuropsych / imaging domain(s). Use ICA method (Calhoun)?

General issues with regards to combining patient data:
Prodromal (clinical high risk) data:
Prodromes may have a different genetic loading. Exclude, or include with covariate for diagnosis?

- Familial/genetic high risk samples – Pittsburgh/Keshavan, DeLisi, Edinburgh/Lawrie & McIntosh, others?
- Clinical high risk – CIDAR (one-third prodromal), others?

Age range in sample:
Outliers for age can significantly influence neuropsychological analyses. Age cut-off for inclusion, or age as a covariate?
Proposal: Adjust raw neuropsych scores for age and standardize across samples. No exclusion of cases.

Dealing with medication use in analyses:
Antipsychotics are a major confound for neuropsych and imaging. Different sites have acquired data on medication use differently. Standardize chlorpromazine equivalence using one specific table.
Should groups standardize and control for these issues within their sample, and send the results, or should antipsychotic use be corrected for across samples?

General issues with regards to combining imaging data:

- Process all scans using same algorithm? e.g., ENIGMA approach using FreeSurfer or FSL. Jessica Turner has done subcortical volume analyses according to the ENIGMA consortium protocols. Perhaps this is a place to start.
- Better to perform genetic association analyses of each sample separately and meta-analyze, however still important to make sure consistency in variables across samples.
- Several people have experience with multisite structural MRI data (René Kahn, Jessica Turner, Jorge Jovicich). DTI data has other issues than sMRI. Marty Shenton's group is developing multisite methods for DTI.

- When one sample is acquired across multiple scanners, use scanner as covariate, and use balanced numbers of cases and controls on individual scanners.

General issues with regards to combining neuropsychological data:
- Proposal to use Timi Toulopoulou's approach to merging cognitive data acquired using different tests to standardize them to be on the same scale (i.e., normalize across the samples using the controls' data).
- Reduce number of variables through factor analysis across all samples. However missing data is a problem so must have all variables in all samples. Could use domains instead. Will depend on how many samples have a certain task.

Additional samples/subjects to consider including:
- COGS multi-site data. Already genotyped on custom SNP chip, not GWAS? Larry Seidman to speak with David Braff
- NAPLS study ~770 clinical high-risk cases (~75 converters); ~275 controls. Contact person: Larry Seidman
- MGH data: Josh Roffman indicated more subjects are available now, but they do not have GWAS.
- DeLisi genetic high risk sample

If sample does not have GWAS, we could genotype the top ~100 SNPs from the SZ risk GWAS meta-analysis. Funding available through Tracey's RO1 grant. May be possible to GWAS genotype some samples.

Data sharing:
Will raw imaging data be sent to a central site for analysis or will each site conduct their own analyses and send results/output to central site?
Informed consent: Some groups do not have consent from all subjects to share data with other groups.
Any data sent to central site should be de-identified and in a standardized format (where possible) to avoid errors or misinterpretation.

***Next step*:**
Clarifying what data is available for each sample - filling in Excel tables

Suggestions will be solicited for information to add to the current lists below. Timi Toulopoulou will look for the information she requested for her published multisite neuropsych study.

Inquire whether each imaging group could send raw data from 10 scans (5 patients, 5 controls) for Jorge Jovicich (consultant) to look at potential issues with data quality.

- General information required from each group/site:
  - Preferred sample name and/or acronym
  - PI(s) and (if relevant) contact person
  - Origin/Country
  - Willing/Able to send de-identified phenotype and genotype data to central site; yes/no

- Clinical and demographic information
  - # cases and diagnostic breakdown (# scz, sczaff-depressed or –bipolar, sczphreniform, etc.)
  - # or proportion within disease stage (genetic and/or clinical high risk, first episode, chronic)
  - # controls, matched to cases?
  - unrelated individuals or family sample
  - # males / # females or male/female ratio
  - ethnicity proportion (e.g., Caucasian, Asian, African American, Hispanic, etc)
  - age range of cases and controls
  - disease history (e.g., how initial diagnosis made (SCID?), duration, remitting or continuous trajectory, degree of functioning (GAF?), positive/negative symptomatology (PANSS?))
  - medication history (e.g., current and lifetime, length of exposure to antipsychotics, first/second generation, treatment responsive/non-responsive, treatment adherence)
  - family history (e.g. FIGS)
  - depression rating, mania rating, anxiety rating (scales used)
  - substance use/abuse (e.g. cannabis use times per week)
  - other clinical/demographic information available (e.g. education level, parental SES)
- Genetic data available
  - # cases, controls with GWAS data from which array (chip version)
  - Any imputation already done, and if so, to which reference panel
- Cognitive data available
  - Was full/partial MATRICS administered
  - List of tests administered across all cognitive domains, test version (e.g. III-revised), country/language version (e.g. US, British, German)
  - Raw scores available?
  - Which standard scores available? (standard scores relative to which reference sample?)
  - Do any studies other than CIDAR have olfaction data (e.g. UPSIT) available?
- Neuroimaging data available and scan parameters
  - scanner type(s) (vendor, model), field strength, head coil, scanner software version(s)
  - sequence types (T1-weighted MP-RAGE, DTI)
  - slice acquisition direction, spatial resolution, # slices, slice gap, slice orientation
  - TR, TE, TI, Flip angle
  - Acquisition duration (Parallel imaging?)
  - For DTI: # b0 volumes, b-values, # gradient directions, # averages
  - # subjects scanned on the same scanner, scanned with the same software version, and with the same protocol.
- Neuroimaging analyses/processing already carried out
  - Variables: e.g., Intracranial volume, Total brain volume, grey matter volume, white matter volume, GM volume ROI measures, Cortical thickness ROI measures, FA tract averages, Total FA average
  - Software (incl. version) used to date: e.g., FSL, Freesurfer, In-house

Gabriëlla Blokland will e-mail the groups to get this information together.

Working groups:
Proposal to initiate working groups around particular issues, e.g.:

- Neuropsychology
- Imaging

Conference calls every other month for Neuropsychology/Imaging.
Some issues (e.g., calculating medication) will overlap across working groups and will need to be discussed with everyone.

Timeline:
4 year RO1 grant; we are at 6 months now.
Neuropsychological data are more easily analyzable first up.
Aim to get neuropsychological data organized by October.

Next meeting:
Proposal for next meeting at World Congress on Psychiatric Genetics to be held in Boston, MA in October.