ENVIRONMENTAL-RISK (E-RISK) LONGITUDINAL TWIN STUDY
CONCEPT PAPER FORM

Proposing Author: Jessica Agnew-Blais

Author's affiliation, phone, and e-mail address: SGDP, King's College London, Jessica.agnew-blais@kcl.ac.uk

Sponsoring Investigator (if the proposing author is a student, a post-doc or a colleague): Louise Arseneault

Proposed co-authors: Louise Arseneault, Andrea Danese, Temi Moffitt, Avshalom Caspi, Benjamin Williams, Karen Sugden, Jasmin Wertz, Dan Belsky, Guilherme Polanczyk, Cathryn Lewis

Provisional Paper Title: ADHD polygenic risk scores and the course of ADHD across development

Date: April 4th, 2018

Objective of the study and its significance:

ADHD is highly heritable and associated with multiple genetic and environmental factors that operate in concert to cause the disorder. A recently published genome-wide association study (GWAS) from the Psychiatric Genomics Consortium (PGC) has identified 12 significant genetic variants associated with ADHD (Demontis 2017). One way to quantify the overall impact of individual single nucleotide polymorphisms (SNPs) associated with ADHD is using polygenic risk scores (PRS), in which SNPs are combined to create an overall genetic liability score. ADHD PRS derived from earlier GWAS has been found to be associated with ADHD symptoms in the general population (Groen-Blokhuis, Middeldorp et al. 2014), conduct disorder among children with ADHD (Hamshere, Stergiakouli et al. 2013), and IQ and working memory (Martin, Hamshere et al. 2015).

Several recent population-based studies, including E-Risk, have identified individuals with late-onset ADHD, who meet criteria for ADHD in adulthood, but not in childhood (Moffitt TE, Houts R et al. 2015, Agnew-Blais, Polanczyk et al. 2016, Caye, Rocha et al. 2016, Riglin, Collishaw et al. 2016). In our paper from E-Risk on late-onset ADHD, we suggested three potential explanations for the late emergence of the disorder. Using a PRS created from the most recent childhood ADHD GWAS we can begin to disentangle these different potential explanations. First, it could be that individuals with late-onset ADHD would have had childhood ADHD, but the disorder was masked in childhood; if this is the case, late-onset individuals should have similar ADHD PRS compared to those who exhibited childhood ADHD. Second, late-onset ADHD could be accounted for by other disorders that present with ADHD-like symptoms, such as depression or alcohol dependence; in this case, the late-onset ADHD group could show elevated PRS scores for these other disorders rather than for childhood ADHD. Third, late-onset ADHD could be a distinct form of ADHD; in this case, late-onset ADHD may not be associated with PRS for childhood ADHD or these other disorders, suggesting a different genetic or environmental aetiology.

Study aims:
Aim 1: ADHD PRS and associated factors
1a. To validate whether ADHD PRS is associated with factors such as childhood ADHD from age 5-12, continuously measured attention problems and hyperactivity symptoms as reported by mums and teachers across ages 5-12, oppositional-defiant disorder and conduct disorder between ages 5-12, and IQ at age 5 in the E-Risk study.

Aim 2: ADHD PRS and ADHD across development

2a. To examine whether ADHD PRS is associated with a more persistent course of ADHD across development. More specifically to examine whether:

(1) ADHD PRS is associated with more frequent ADHD diagnoses across childhood (at ages 5, 7, 10 and 12)
(2) ADHD PRS is associated with a persistent course of ADHD from childhood to young adulthood, compared with a remitting course of the disorder.

2b. To investigate the genetic underpinnings of late-onset ADHD by examining whether ADHD PRS is significantly predictive of late-onset ADHD, suggesting childhood and late-onset ADHD share genetic aetiology, or late-onset ADHD is associated with PRS for other mental health disorders (depression, anxiety, alcohol use), suggesting other psychopathology may account for late-onset ADHD. We will also examine how the PRS for depression, anxiety and alcohol use in the late-onset group compare to groups who meet criteria for these disorders at age 18.

Statistical analyses:

Aim 1. First, I will calculate an ADHD PRS for each participant with SNPs identified by the most recent ADHD GWAS (Demontis 2017) using the PRSice2 software. Given that my fellowship focuses on skills development in statistical genetics, creating the ADHD PRS from scratch is an important part of my training and why I have requested individual participant genotypes below. The PRS will be constructed not just from genome-wide significant SNPs identified by GWAS, but using different statistical significance cut-offs, in line with approaches operationalized in the PRSice2 software. ADHD PRS scores will be used in regression analyses, with ancestry-informative principal component covariates, to determine whether ADHD PRS is associated with having ADHD in childhood, ADHD symptoms levels in the entire study population, and associated factors such as childhood ODD, CD, and IQ. The ADHD PRS I will create here at the IoPPN will be compared to the ADHD PRS created at Duke as a validation step and for the sake of consistency.

Aim 2.

2a. I will assess whether ADHD PRS is associated with persistence vs remission of ADHD by age 18 using logistic regression adjusting for ancestry-informative principal component covariates.

2b. I will assess whether ADHD PRS is associated with late-onset ADHD using regression analyses to compare the mean ADHD PRS in the late-onset group to the mean among non-ADHD controls, and those with persistent and remitted ADHD. I will also calculate the PRS for depression, anxiety and alcohol use to determine whether those with late-onset ADHD have higher mean PRS for these other disorders/behaviors compared to non-ADHD controls and to those with persistent and remitted ADHD. There is variation in the amount of GWAS data available for creation of the PRS for the different disorders we have speculated may be related to late-onset ADHD. There are large GWAS available for depression, and the PGC and other consortia have posted online genome-wide statistics of GWAS for MDD and anxiety disorders. Additionally, the UK Biobank includes information on related phenotypes including alcohol and cannabis use (http://biobank.ctsu.ox.ac.uk/crystal/label.cgi?id=136). Working with the statistical genetics group here at King's we would be able to generate GWAS information using UK Biobank for disorders where currently publically available summary statistics are not published.

In order to understand whether results regarding ADHD PRS and late-onset ADHD replicate across cohorts, I will apply to conduct equivalent analyses in the Dunedin Study in New Zealand and the Pelotas Study in Brazil (genotyping of a subsample of the 1993 Pelotas cohort will be done using the additional funding I've received from King's). Working with the Pelotas Cohort has the added interest of adding to our understanding of how PRS largely created in GWAS of individuals with European ancestry perform in a more ancestrally mixed population. The paper from the Dunedin cohort looking at adult-onset ADHD found ADHD PRS was not associated with adult-onset ADHD, but we can update these analyses with an ADHD PRS generated based on the most recent and largest ADHD GWAS.

Variables Needed at Which Ages (names and labels):

Study: E-Risk

Age 5:
FAMILYID ID Family
ATWINID ID Twin 1
BTWINID ID Twin 2
SAMPSEX Sex of twins
ZYGOSITY Zygosity of twins
RORDERP5 Random order variable

SESWQ35 Social Class Composite
IQE5 Pro-rated IQ score

TDEPMM5 DSM-IV Depression - After Twins Birth (incl disability criteria)

Age 7:
CHAOSM7—Choas in the home

Age 12:
ANYCDDX_EMT512 Any CD dx from 5 to 12, mum/tchr, Elder
ANYODDDX_EMT12 Any ODD dx from 5 to 12, mum/tcher, Elder [I think this variable still needs to be made]

'Inattention symptom count - (M & T sum, averaged across ages 5, 7, 10, 12) – ET
'Hyperactive/Impulsive symptom count - (M & T sum, averaged across ages 5, 7, 10, 12) – ET
'Total Inattentive/Hyperactive/Impulsive symptom count - (M & T sum, averaged across ages 5, 7, 10, 12) – ET

HARME512 Child Harm Phase 5-12 - Elder Twin
POLYVE512C Extent of Polyvictim(Truncated @3), 5-12, E-Twin

Age 18:
DXADHD5X_18E DSM-5 ADHD Dx [incl 4 NEET & meds] – P18—ET
ADHD4CATE18 ADHD group status

INF_ADHD18E # any informant adhd symptoms, max=7, 18 e-twin

SR_INSUM18E # DSM-5 Inattn symp, Max=9, 18, E-Twin
SR_HYSUM18E # DSM-5 Hyper/Imp symp, Max=9, 18, E-Twin
SR_SYMTOT18E DSM-5 Inattn/Hyper/Imp Symp, Max=18, 18 E-Twin

DXMDEE18 Major depressive episode, dsm4_base – P18 – Elder
DXGADE18 Gen Anxiety Disorder, dsm4_based – P18 – Elder
DXALCDEPE18 Alcohol dependent, dsm4_based - P18 - Elder
DXMARJ5E18 Marijuana use disorder, dsm5 – P18 - Elder

Participant Genotypes
Ancestry principal components

References cited:

Agnew-Blais, J. C., et al. (2016). "Evaluation of the persistence, remission and emergence of ADHD in young adulthood: Results from a longitudinal, prospective population-based cohort." JAMA Psychiatry **73**(7): 713-720.

Belsky, D. W., et al. (2014). "Gene-environment interaciton research in psychiatric epidemiology: A framework and implications for study design." Soc Psychiatry Psychiatr Epidemiol **49**(10): 1525-1529.

Caye, A., et al. (2016). "Does the adult ADHD trajectory always begin with childhood ADHD? Negative evidence from a large birth cohort." JAMA Psychiatry **73**(7): 705-712.

Caye, A., et al. (under review). "A risk score calculated in childhood to predict adult Attention-deficit/Hyperactivity disorder: building the path to individualized prevention in mental health."

Demontis, e. a. (2017). "Discovery of the first genome-wide significant risk loci for ADHD."

Groen-Blokhuis, M. M., et al. (2014). "Attention-deficit/hyperactivity disorder polygenic risk scores predict attention poblems in a population-based sample of children." J Am Acad Child Adol Psychiatry **53**(10): 1123-1129.

Hamshere, M. L., et al. (2013). "Shared polygenic contribution between childhood attention-deficit disorder and adult schizophrenia." Br J Psychiatry **203**(2): 107-111.

Martin, J., et al. (2015). "Neurocognitive abilities in the general population and coposite genetic risk scores for attention-deficit hyperactivity disorder." J Child Psychol Psychiatr **56**(6): 648-656.

Moffitt TE, et al. (2015). "Is adult ADHD a childhood-onset neurodevelopmental disorder? Evidence from a four-decade longitudinal cohort study." Am J Psychiatry **172**(10): 976-977.

Riglin, L., et al. (2016). "Association of Genetic Risk Variants With Attention-Deficit/Hyperactivity Disorder Trajectories in the General Population." JAMA Psychiatry **73**(12): 1285-1292.

Vernon-Feagans, L., et al. (2016). "Predictors of behavioral regulation in kindergarten: Household chaos, parenting, and early executive functions." Devel Psychopathol **52**(3): 430-4441.

Data Security Agreement

Provisional Paper Title	ADHD polygenic risk scores and the course of ADHD across development
Proposing Author	Jessica Agnew-Blais
Today's Date	April 4th 2018

Please keep one copy for your records
(Please initial your agreement)

_jab___I am familiar with the King's College London research ethics guidelines (https://www.kcl.ac.uk/innovation/research/support/ethics/about/index.aspx) and the MRC good research practice guidelines (https://www.mrc.ac.uk/research/policies-and-guidance-for-researchers/good-research-practice/)

__jab__My project has ethical approval from my institution.

_jab___My computer is (a) encrypted at the hard drive level, (b) password-protected, (c) configured to lock after 15 minutes of inactivity, AND (d) has an antivirus client which is updated regularly.

_jab___I will treat all data as "restricted" and store in a secure fashion.

_jab __I will not share the data with anyone, including students or other collaborators not specifically listed on this concept paper.

_ jab _I will not merge data from different files or sources, except where explicit approval has been given by the PI.

_ jab _ I will not post data online or submit the data file to a journal for them to post.
Some journals are now requesting the data file as part of the manuscript submission process. The E-Risk Study cannot be shared because the Study Members have not given informed consent for unrestricted open access. Speak to the study PI for strategies for dealing with data sharing requests from Journals.

_ jab _Before submitting my paper to a journal, I will submit my draft manuscript and scripts for data checking, and my draft manuscript for co-author mock review, allowing three weeks.

_ jab __I will submit analysis scripts and new variable documentation to project data manager after the manuscript gets accepted for publication.

____ **For projects using location data:** I will ensure geographical location information, including postcodes or geographical coordinates for the E-Risk study member's homes or schools, is never combined or stored with any other E-Risk data (family or twin-level data)

__ jab __**For projects using genomic data:** I will only use the SNP and/or 450K data in conjunction with the phenotypes that have been approved for use in this project at the concept paper stage.

Signature: .....Jessica Agnew-Blais.....