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

Proposing Author: Susanna Roberts

Author's affiliation, phone, and e-mail address:
x0837: SGDP Centre, IoPPN, King's College London; susanna.roberts@kcl.ac.uk

Sponsoring Investigator (if the proposing author is a student, a post-doc or a colleague):
Helen Fisher & Jon Mill

Proposed co-authors:
Louise Arseneault, Avshalom Caspi, Chloe Wong, Temi Moffitt, Eilis Hannon, Sarah Marzi, Emma Dempster

Provisional Paper Title:
DNA methylation and psychotic experiences at age 18

Date:
15th June 2017

Objective of the study and its significance:

Psychotic experiences in adolescence (e.g., hearing or seeing things that others don't, having unusual beliefs, etc.) are less common than in childhood, but are predictive of both psychotic disorders and other psychopathologies in adulthood [1], and the clinical relevance of psychotic phenomena increases later in adolescence [2]. Therefore, uncovering biomarkers associated with pre-clinical psychotic experiences during this period has the potential to facilitate early identification of individuals at-risk of a range of mental health problems, and improve targeting of preventive interventions.

Despite high estimates of genetic influence on psychotic experiences, there is still considerable discordance within monozygotic (MZ) twin-pairs, indicating that person-specific non-genetic factors are also important in mediating their onset [3]. Epigenetic processes are dynamic mechanisms that have the potential to regulate gene expression without changing the underlying genetic sequence, and as such may be one potential biomarker in the aetiology of psychotic experiences. Research in MZ twin pairs discordant for psychosis and in patient-control samples has identified DNA methylation differences associated with clinically-relevant psychosis [4, 5]. Previous work in E-Risk found epigenetic variation (at age 10) was associated with childhood psychotic symptoms (at age 12) [6]. However, to date, no studies have investigated epigenetic variation with respect to psychotic phenomena in late adolescence.

In this project, we plan to investigate whether DNA methylation patterns at age 18 are associated with concurrent psychotic experiences in the full E-Risk sample and then in discordant monozygotic twins. The latter represents a powerful strategy in epigenetic epidemiology as identical twins are matched for a number of potential confounding effects such as age, sex, genetics, early family environment, and family psychiatric history [7].We also propose to use the E-Risk sample to attempt to replicate the top findings from an analysis currently being conducted in the ALSPAC sample.

Statistical analyses:

Full project
DNA methylation has been quantified from whole blood samples in E-Risk at age 18 using the Illumina Infinium HumanMethylation 450K array, giving quantitative data for >480,000 CpG sites across the genome. Blood cell counts will be quantified using the Houseman algorithm [1], and the methylation data adjusted for cell composition and batch using linear regression. Results will be controlled for multiple-testing using the FDR. We propose to investigate the association between age-18 psychotic experiences

and DNA methylation at age 18 in the full sample. We will then continue by comparing this data in twins discordant for concurrent psychotic experiences (i.e. at age 18). We will also conduct analyses controlling for smoking, given its important role as a confounder in methylation data, particularly in the context of mental health outcomes.

Main analyses

1. We will test the association between DNA methylation and psychotic experiences at age 18 in the full sample.
    a. Linear regression will be used to test for the association between DNA methylation and i) any psychotic experiences, and ii) the extent of psychotic experiences (ordinal scale) at age 18, controlling for sex, zygosity, cell composition, batch and the non-independence of twin observations.
    b. For the top probes identified, models will be rerun in subgroups with psychotic symptoms at age 12, to compare those with and without psychotic experiences at age 18 (to test for associations with persistence of psychotic phenomena).
2. We will test the association between DNA methylation and psychotic experiences at age 18 in MZ and DZ twin pairs discordant for psychotic experiences at age 18 and then restricted to only discordant MZ twin pairs.
    a. Paired t-tests will be used to test for differences in DNA methylation between psychotic experience-discordant twins.
3. Models will be repeated controlling for smoking.
    a. Recorded smoking status will be compared to smoking status derived from measured DNA methylation at known smoking-related CpG sites.
    b. Linear regression will be used to adjust the data for the effects of smoking (pack years) and analyses will be re-run using the adjusted data.
4. Gene ontology enrichment analyses will be conducted to identify biological processes overrepresented in the top results of each analysis.
5. We will also explore whether the top probes identified by Fisher et al [6] as associated with psychotic symptoms at age 12 (in age 10 DNA methylation data) show a consistent direction of effect in the age-18 data.

Exploratory analysis

6. Numbers permitting, analyses will be repeated using clinically-verified psychotic symptoms at age 18 as the outcome measure to test associations using a more conservative phenotype.

Project replication

7. We aim to check any DNA methylation markers associated with psychotic experiences in the E-Risk sample in (i) an independent sample of adult post-mortem brain tissue (prefrontal cortex) from 20 schizophrenia cases and 23 non-psychiatric controls archived in the London Brain Bank for Neurodegenerative Disorders and from 18 schizophrenia cases' and 15 non-psychiatric controls' brains obtained from the Douglas Bell-Canada Brain Bank, Montreal [3], and (ii) in an independent sample of schizophrenia discordant MZ twins using data held by Emma Dempster.

Replication for ALSPAC paper

We also plan to replicate the top sites from a study we have conducted in ALSPAC looking at DNA methylation in cord blood, and samples at ages 7 and 15-17 and the trajectory of psychotic symptoms. Probes of interest were identified using a three-stage analysis plan: time-point specific EWAS; investigation of the top probes at each time-point with psychotic symptom trajectories; and longitudinal analysis of the top probes across development. For this replication, we will focus on the top probes detected in adolescent DNA methylation data (age 15-17) as these are likely to be the most closely aligned with the available data in E-Risk.

8. To attempt to replicate the top probes from these analyses in E-Risk, we will:
    a. Use linear regression to test the association between DNA methylation at the probes of interest and psychotic experiences at age 18 in the whole sample, controlling for sex, zygosity, batch, cell-type proportion estimates, and smoking, and including robust standard errors to account for family structure.
    b. Numbers permitting, analyses will be repeated using clinically verified psychotic symptoms as the outcome measure.

Variables Needed at Which Ages (names and labels):

Study: E-Risk

Core
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)

Age 12
PSYSYMP01E12 (Psychosis symptom count, verified coding; 0, 1+)

Age 18
PSYEXPE18 (Psychotic experiences scale elder)
PSYEXPCE18 (Psychotic experiences (cat) elder)
PSYSYMP01E18 (Psychosis symptom count; 0, 1+, elder)
PSYEXPY18 (Psychotic experiences scale younger)
PSYEXPCY18 (Psychotic experiences (cat) younger)
PSYSYMP01Y18 (Psychosis symptom count; 0, 1+, younger)

SMKPKYRE18 (Smoking - pack years, ages 12 to 18)
SMKCNUME18 (Smoking - current number of cigarettes - *Number of cigarettes smoked per day at age 18 or age 19 smoking level if that age at interview)*

Illumina 450k DNA methylation data from peripheral blood at age 18
+ related variables (e.g., batch number)

References cited:

1. Kelleher, I., et al., *Prevalence of psychotic symptoms in childhood and adolescence: a systematic review and meta-analysis of population-based studies.* Psychological Medicine, 2012. **42**(9): p. 1857-1863.

2. Kelleher, I., et al., *Clinicopathological significance of psychotic experiences in non-psychotic young people: evidence from four population-based studies.* British Journal of Psychiatry, 2012. **201**(1): p. 26-32.

3. Zavos, H.M.S., et al., *Consistent etiology of severe, frequent psychotic experiences and milder, less frequent manifestations: a twin study of specific psychotic experiences in adolescence.* JAMA psychiatry, 2014. **71**(9): p. 1049-1057.

4. Dempster, E.L., et al., *Disease-associated epigenetic changes in monozygotic twins discordant for schizophrenia and bipolar disorder.* Human molecular genetics, 2011. **20**(24): p. 4786-4796.

5. Pidsley, R., et al., *Methylomic profiling of human brain tissue supports a neurodevelopmental origin for schizophrenia.* Genome biology, 2014. **15**(10): p. 483.

6. Fisher, H.L., et al., *Methylomic analysis of monozygotic twins discordant for childhood psychotic symptoms.* Epigenetics, 2015. **10**(11): p. 1014-1023.

7. Mill, J. and B.T. Heijmans, *From promises to practical strategies in epigenetic epidemiology.* Nature reviews. Genetics, 2013. **14**(8): p. 585-594.

8. Houseman, E.A., et al., *DNA methylation arrays as surrogate measures of cell mixture distribution.* BMC Bioinformatics, 2012. **13**(1): p. 86.

Data Security Agreement

Provisional Paper Title	DNA methylation and psychotic experiences at age 18
Proposing Author	Susanna Roberts
Today's Date	14/06/17

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

__SR_ I am current on Human Subjects Training (CITI (www.citiprogram.org) or training in human subject protection through my post or courses.

_SR My project is covered by Duke or King's IRB OR I have /will obtain IRB approval from my home institution.

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

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

__SR_ 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 Terrie or Avshalom for strategies for dealing with data sharing requests from Journals.

__SR_ 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.

__SR I will submit analysis scripts and new variable documentation to project data manager after manuscript gets accepted for publication.

SR I will return all data files to the Data Manager after the project is complete. Collaborators and graduates of DPPP may not take a data file away from the DPPP office. The data remains the property of the Study and cannot be used for further analyses without express, written permission.

SR 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)

Signature: ...  .....

CONCEPT PAPER RESPONSE FORM

A. To be completed by the proposing author

Proposing Author:

X I have read the E-Risk data-sharing policy guidelines and agree to follow them

Provisional Paper Title: DNA methylation and psychotic experiences at age 18

Potential co-authors: Helen Fisher, Jon Mill, Louise Arseneault, Avshalom Caspi, Chloe Wong, Temi Moffitt, Eilis Hannon, Sarah Marzi, Emma Dempster

Potential Journals:

Intended Submission Date (month/year): November 2017

Please keep one copy for your records and return one to Louise (louise.arseneault@kcl.ac.uk)

B. To be completed by potential co-authors:

☐ X Approved ☐ Not Approved ☐ Let's discuss, I have concerns

Comments:

Please check your contribution(s) for authorship:

☐ x Conceptualizing and designing the longitudinal study

☐ Conceptualizing and collecting one or more variables

☐ Data collection

☐ Conceptualizing and designing this specific paper project

☐ Statistical analyses

☐ x Writing

☐ x Reviewing manuscript drafts

☐ x Final approval before submission for publication

☐ Acknowledgment only, I will not be a co-author

Signature: .....