Concept Paper

Provisional Paper Title:
Epigenome-wide Association Study of Attention problems

Proposing Authors:
From Duke University:
Terrie Moffitt, Karen Sugden, Avshalom Caspi,

From VU Amsterdam:
Jenny van Dongen, j.van.dongen@vu.nl,
Nuno R. Zilhão,
Dorret Boomsma di.boomsma@vu.nl
Vrije Universiteit, Dept Biological Psychology, Netherlands Twin Register
Van der Boechorststraat 1, 1081 BT, Amsterdam, The Netherlands
Room 2B41
phone: 31-20-5988787 / fax: 31-20-5988832
email: di.boomsma@vu.nl
twitter: @NTRscience
url: www.tweelingenregister.org

From Radboudumc, Nijmegen
Barbara Franke, Barbara.Franke@radboudumc.nl
Department of Human Genetics (855) and Department of Psychiatry
Donders Institute for Brain, Cognition and Behaviour, Centre for Medical Neuroscience
Radboud university medical center
PO Box 9101, 6500 HB Nijmegen, The Netherlands
Tel: +31-24-3610181
E-mail barbara.franke@radboudumc.nl
URL: http://www.ru.nl/people/donders/franke-b/

Author's Email: tem11@duke.edu,

P.I. Sponsor:
(if the proposing author is a student or colleague of an original PI)

Today's Date: 1 august, 2017

Please describe your proposal in 2-3 pages with sufficient detail for helpful review.

Objective of the study:

NOTE: Data analysis for this project will take place at Duke University.

Colleagues at the Vrije Universiteit Amsterdam conducted an epigenome-wide association study (EWAS) of attention problems in adolescents and adults assessed with the CAARS ADHD index. The Illumina 450k methylation data from whole blood were collected by the Netherlands Twin Register (NTR) biobank in collaboration with BBMRI-BIOS. This study identified several differentially methylated positions (CpGs) at genome-wide or nominal

significance that they would like to replicate in independent cohorts. Therefore, they would like to ask the E-risk study and the Dunedin Study to perform a look-up analysis of the top CpGs (N=29), and ADHD symptoms (total symptoms, hyperactivity/impulsivity and inattention) assessed in adults with a measure as similar as possible to the CAARS and share summary statistics. Ideally, the researchers would like to also receive genome-wide summary statistics to allow estimating and adjusting for genome-wide test statistics bias and inflation. However, the primary interest is to replicate the 29 top sites. A power analysis was performed for the look-up (alpha= 0.0017, i.e. 0.05/29), which indicated that with a replication cohort of > 700 subjects and given the effect sizes observed in NTR, there would be decent power for at least a few top sites.

The following 12 items from the Conners Adult ADHD Rating Scales (CAARS)* ADHD index were used to score ADHD symptoms in the NTR cohort.

1. CAARS-4 item 3: I am always on the go as if driven by a motor,
2. CAARS-4 item 5: I have a short fuse/hot temper,
3. CAARS-4 item 7: I still throw tantrums,
4. CAARS-4 item 10: I avoid new challenges because I lack faith in my abilities,
5. CAARS-4 item 11: I feel restless inside even if I sit still,
6. CAARS-4 item 12: Things I hear or see distract me from what I am doing,
7. CAARS-4 item 15: I am an underachiever,
8. CAARS-4 item 17: I cannot get things done unless there is an absolute deadline,
9. CAARS-4 item 20: I intrude on others activities,
10. CAARS-4 item 23: Sometimes my attention narrows so much that I'm oblivious to everything else; other times it's so broad that everything distracts me,
11. CAARS-4 item 24: I cannot keep my mind on something unless it's really interesting,
12. CAARS-4 item 28: My past failures make it hard for me to believe in myself,

Answer categories (1=never, 2=once in a while, 3=often, 4=frequently; items were summed

* Conners, CK; Erhardt, D; Sparrow, E. *Conners' Adult ADHD Rating Scales (CAARS) technical manual*. (North Tonawanda, NY; Multi-Health Systems, 1999).

At Duke, the first thing we propose to do is undertake an EWAS of an appropriate ADHD symptoms scale [z1]in the Dunedin Study (by appropriate we mean the best measure, with the best distribution, and one that is age-matched to the blood collection). This will be the Self-reported ADHD symptoms scale at Phase 38.

The second thing we propose to do is undertake an EWAS of appropriate an ADHD symptoms scale in the E-risk Study. This will be: Self-reported ADHD symptoms scale at Phase 18

Data analysis methods:

Model:

For all methylation sites, please run:
Model 1: Methylation ~ ADHD + Age + Sex + Smoking + WBC percentages + Technical + Cohort Specific
Model 2: Methylation ~ ADHD + Age + Sex + Smoking + WBC percentages + Technical + Cohort Specific + genetic principal components

* ADHD: Ideally, we like to run these models for 3 ADHD phenotypes: 1=total ADHD symptoms, 2=inattention symptoms and 3=hyperactivity/impulsivity symptoms.

Note: If the cohort includes related individuals (e.g. family members, twins), please apply a statistical approach that takes the clustering of data into account (e.g. gee or linear mixed models).

Covariates
Age= Age when DNA sample was collected
WBC= White blood cell percentage in the same blood sample from which DNA was extracted. If you did not measure white blood percentages in the same sample as used for the DNA methylation measurement, please estimate WBC percentages using a prediction method (e.g. Houseman's reference based method). For computational reasons, please do not include multiple WBC that are highly correlated with each other or that show very little variation between people in your cohort. For example, in the NTR, we use the following WBC as covariates: monocyte percentage, eosinophil percentage, neutrophil percentage.
Technical covariates + Cohort Specific covariates Please correct for technical (batch) covariates and other cohort-specific covariates as you deem necessary. For example, at the NTR, we include 450k array row and either sample plate or (carefully chosen) principal components from the methylation data.
Smoking= Smoking status at the moment of blood sampling, 3 levels: 0=never smoked, 1=former smoker, 2=current smoker.

Genetic PCs, Model 2: The second model accounts for population ancestry. In the NTR, we excluded a small group of 'ethnic outliers' (individuals with a mixed genetic background, identified based on genetic PCs) to select only individuals of Dutch ancestry for this model. Secondly, we included 3 principle components based on genome-wide SNP data that reflect genetic ancestry in the Dutch population as covariates in this model.

Software and example R-script
Feel free to use your own analysis pipeline or preferred software to run the models outlined above. Example R-code with instructions is provided in:
Example LM script_012016.r [suited for cohorts that include unrelated subjects]
Example gee script_012016.r [suited for cohorts that include related subjects (family members)]

Variables needed at which ages:

Dunedin Study: Genome-wide methylation from Phase 38 blood.
Self-reported ADHD symptoms at Phase 38, described in:
Moffitt TE, Houts R, Asherson P, Belsky, DW, Corcoran, DL, Hammerle M, Harrington HL, Hogan S, Meier M, Polanczyk GV, Poulton R, Ramrakha S, Rohde LA, Sugden, K, Williams, B, Caspi A (2015). Is adult ADHD a childhood-onset neurodevelopmental disorder? A 4-decade longitudinal cohort study. American J of Psychiatry. 172, 967-977.

E-Risk: Genome-wide methylation from Phase 18 blood.
Self-reported ADHD symptoms at Phase 18, described in: Agnew-Blaise, J, Polanczyk, G, Danese, A, Wertz, J, Moffitt, TE, Arseneault, L (2016). Persistence, remission and emergence of ADHD in young adulthood: Results from a longitudinal, prospective population-based cohort. JAMA-Psychiatry..

Covariates for both cohorts:
Sex,
Age at blood draw (not needed, as all participants in each cohort were the same age)
WBC= White blood cell percentage in the same blood sample
Smoking status at time of blood draw
Genotype PCs

Significance: A better understanding of the genetics of ADHD symptoms in the population and (knowing about the phenotypic and genetic continuum between population symptom scores and clinical disorder) hypotheses to be tested for clinical ADHD.

Data Security Agreement

A

Provisional Paper Title	**Epigenome-wide Association Study of Attention problems**
Proposing Authors	Moffitt, Sugden, Caspi
Today's Date	August 1, 2017

Please keep one copy for your records and return one to the PI Sponsor

Please initial your agreement

x	I am current on Human Subjects Training (CITI (www.citiprogram.org) or equivalent)
x	My project is covered by Duke or Otago IRB OR I have /will obtain IRB approval from my home institution.
x	I will treat all data as "restricted" and store in a secure fashion.
x	I will not share the data with anyone, including students or other collaborators not specifically listed on this concept paper.
x	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 Dunedin 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.*
x	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
x	I will submit analysis scripts and new variable documentation to project data manager after manuscript gets accepted for publication.
x	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.

Signature:

CONCEPT PAPER RESPONSE FORM
A

Provisional Paper Title	**Epigenome-wide Association Study of Attention problems**
Proposing Authors	Moffitt, Sugden, Caspi from Duke
Other Contributors	Van Dongen, Zilhão, Boomsma from Vrije Universiteit Amsterdam Barbara Franke, from Radboudumc, Nijmegen Richie Poulton from Otago Univ., Dunedin Study director. Louise Arseneault from King's College, E-Risk director. Jon Mill from Exeter University, EWAS leader for E-Risk.
Potential Journals	To be decided by the Dutch team
Today's Date	August 1 2017
Intended Submission Date	Fall 2017

Please keep one copy for your records and return one to the proposing author

B. To be completed by each potential co-author and returned to the proposing author:

	Approved
	Not Approved
	Let's discuss, I have concerns

Comments:

Please check your contribution(s) for authorship:

	Conceptualizing and designing the longitudinal study
	Conceptualizing and collecting one or more variables
	Data collection
	Conceptualizing and designing this specific paper project
	Statistical analyses
	Writing
	Reviewing manuscript drafts
	Final approval before submission for publication
	Acknowledgment only, I will not be a co-author

Signature:

_____