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

Proposing Author: Sarah Marzi

Author's affiliation, phone, and e-mail address:
MRC Social, Genetic, and Developmental Psychiatry Centre
Institute of Psychiatry, Psychology and Neuroscience
King's College London
London SE5 8AF, United Kingdom

Phone: 02078480748

Email: sarah.marzi@kcl.ac.uk

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

Proposed co-authors: any interested E-Risk investigators, Karen Sugden, Ben Williams

Provisional Paper Title: Epigenetic correlates of individual differences in peripheral inflammation levels among victimised children

Date: 10/02/17

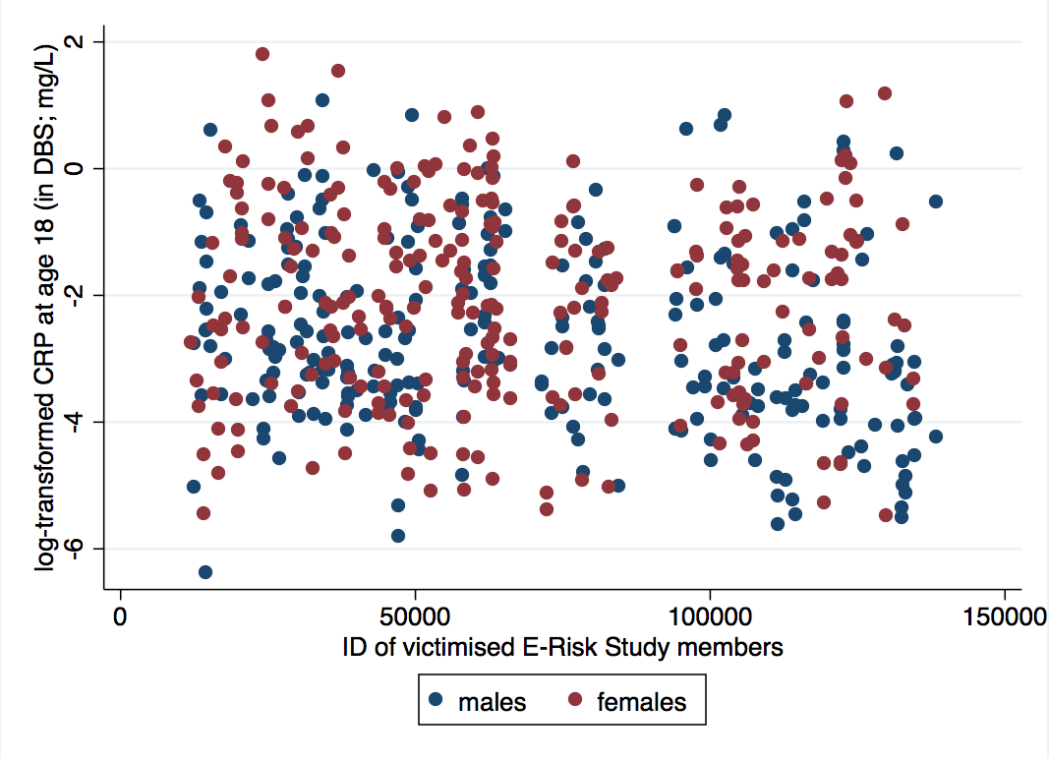
Objective of the study and its significance:

Our team and others have shown that children exposed to violence victimisation are more likely to have high inflammation levels in later life compared to non-victimised peers, possibly leading to greater disease risk in later life (Danese & Baldwin, 2017). However, these group differences fail to capture the remarkable resilience that many victimised children display in the face of adversity. Indeed, there is significant variation in inflammation levels within a group of E-Risk Study members who were victimised as children (see ***Figure*** below). The heterogeneity in inflammation levels among victimised children highlights individual differences in resilience within the at-risk population. We will examine the possible biological correlates of this heterogeneity by testing whether DNA methylation probes predict inflammation levels within a group E-Risk Study members who were victimised as children. The analyses will be based on whole blood DNA from Study members at age 18, which has been used to quantify genome-wide patterns of DNA methylation with the Illumina Infinium HumanMethylation 450K array.

Significance

Understanding what makes some children more resilient and others more vulnerable to the effects of violence victimisation is important to elucidate the biological pathways underlying these effects and to identify targets for preventative and treatment interventions.

***Figure**. Scatterplot of CRP value among E-Risk Study members victimised in childhood (childhood poly-victimisation score >=1; n=591). CRP values are sorted on the X-axis by Study members' ID for illustration purposes*



Statistical analyses:

1. **VALIDATION**: What are the DNA methylation markers associated with CRP levels at age 18 years? We will be able to compare our findings with a recent EWAS (Ligthart, 2016, Genome Biol), although will expect some discrepancies because of the different nature of the specimens (serum versus dried blood spots).
   Method: Linear model regressing DNA methylation onto CRP levels at age 18 in E-Risk Study members, controlling for family relatedness (robust standard errors). Further covariates to be included in all analyses: batch, cell-type proportion, sex.
2. **HYPOTHESIS TESTING**: Within Study members who were victimised in childhood, is DNA methylation in whole blood associated with CRP levels at 18 years?
   Method: Linear model regressing DNA methylation onto CRP levels at age 18 in the subsample of victimized individuals, controlling for family relatedness (robust standard errors). Further covariates to be included in all analyses: batch, cell-type proportion, sex.

3. **SPECIFICITY**: Within Study members who were non-victimised in childhood, which DNA methylation probes predict CRP levels at 18 years? What is the overlap between probe-sets in (2) and (3)?
   Method: Linear model regressing DNA methylation onto CRP levels at age 18 in the subsample of non-victimized individuals, controlling for family relatedness (robust standard errors). Compare effect sizes and p-values at top differentially methylated probes between probe-sets in (2) and (3).
4. **CONFOUNDING**: Is the DNA methylation-based prediction of CRP levels explained by potential confounding factors, such as latent or measured genetic influence (polygenic risk score for CRP), or artefacts linked to sex, cell type distribution, or measurement error in exposure (number of child victimisation types or re-victimisation in adolescence)?
   Methods: a) Linear model regressing DNA methylation onto CRP levels at age 18 in the subsample of victimized individuals, controlling for family relatedness (robust standard errors) and the specific individual confounders.
   b) Twin difference model regression differences in DNA methylation onto differences in CRP, controlling for cell-type proportion differences.
5. **PATHWAYS**: What is the biological function of the identified DNA methylation probe-set - as derived from gene ontology and targeted ontology?
   Method: a) Pathway analysis for gene ontology is based on the set of significant probes, identifying molecular pathways enriched in these, compared to the whole array background. b) Targeted ontology, specifically queries the enrichment of candidate pathways in the set of differentially methylated probes.
6. **DIFFERENTIALLY METHYLATED REGIONS**: Using dimension reduction techniques (comb-P) and modules of co-methylated probes (WGCNA), can we identify regions of differential methylation associated with CRP levels in the subsample of individuals victimized in childhood?
   Method: a) Using comb-P on the results of the CRP EWAS in the victimized subsample to combine P-values into local regions. b) WGCNA will be used to identify networks of co-methylated probes and associate these with CRP levels as described above at the individual probe level.
7. **REPLICATION**: We will seek to identify other samples where we could test similar hypothesis

Variables Needed at Which Ages (names and labels):

Study:

Age 5

- FAMILYID Unique family identifier
- ATWINID Twin A ID (ex chkdg)
- BTWINID Twin B ID (ex chkdg)
- RORDERP5 Random Twin Order
- RISKS Sample Groups
- COHORT Cohort
- SAMPSEX Sex of Twins: In sample
- ZYGOSITY Zygosity
- SESWQ35 Social class composite

Age 12

- POLYVE512 Extent of polyvictim, 5-12, Elder
- POLYVY512 Extent of polyvictim, 5-12, Younger
- CRPEmgl CRP mGl – Elder (Germfighters)
- CRPYmgl CRP mGl – Younger (Germfighters)

Age 18

- FinalconcentrationmgLe18 CRP elder
- FinalconcentrationmgLy18 CRP younger
- WAISTHIPE18 Waist-hip ratio (accounting for pregnant women) - Elder
- WAISTHIPY18 Waist-hip ratio - Younger
- Body temp @ CRP measure elder and younger
- POLYVCTZE18 Poly-victimisation count - Elder
- POLYVCTZY18 Poly-victimisation count – Younger
- Revictimisation (based on latent class analysis by Avshalom)
- EWAS data (including batch, cell types, predicted zygosity and DNA methylation age)
- polygenic risk score for CRP (derived by Karen Sugden/Dan Belsky)
- smoking pack-years

References cited:

Danese A, Baldwin JR. Hidden Wounds? Inflammatory Links Between Childhood Trauma and Psychopathology. Annu Rev Psychol. 2017 Jan 3;68:517-544

Ligthart S, …, Dehghan A. DNA methylation signatures of chronic low-grade inflammation are associated with complex diseases. Genome Biol. 2016 Dec 12;17(1):255.

Data Security Agreement

Provisional Paper Title	EPIGENETIC CORRELATES OF INDIVIDUAL DIFFERENCES IN PERIPHERAL INFLAMMATION LEVELS AMONG VICTIMISED CHILDREN
Proposing Author	Sarah Marzi
Today's Date	10/02/17

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

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

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

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

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

SJM 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.

SJM 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.

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

SJM 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.

SJM I will ensure geographical location information, including postcodes or geographical coordinates for the E-Risk study member's homes or schools, is <u>never</u> combined or stored with any other E-Risk data (family or twin-level data)

Signature: