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

Proposing Author: Rachel Latham

Author's affiliation, phone, and e-mail address: SGDP Centre, IoPPN, King's College London, rachel.latham@kcl.ac.uk

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

Proposed co-authors: Louise Arseneault, Valeria Mondelli, Christian Kieling, Thiago Rocha, Andrea Danese, Line Rasmussen, Brandon Kohrt +other interested E-Risk investigators

Provisional Paper Title: Preliminary investigation of associations between individualised risk at age 12 of future depression and age-18 inflammation in a UK longitudinal cohort study

Date: 25 February 2021

Objective of the study and its significance:

Background

Major depressive disorder (MDD) – a leading contributor to global mental-health related disease burden[1] – commonly develops in adolescence and can have a chronic course throughout life[2]. The early identification of those who are at risk for developing MDD is therefore critical to enable preventative intervention.

Addressing this issue of early identification, Rocha and colleagues[3] recently developed, and then validated in the E-Risk cohort, a multivariable prognostic model to predict individual risk in early adolescence of developing MDD at age 18 for adolescents with no previous depressive symptoms. The E-Risk prediction model used social and demographic risk factors at age 12 (sex, skin colour, drug use, school failure, social isolation, fight involvement, child maltreatment, history of running away from home) to predict the risk of developing MDD at age 18 and did so with reasonable accuracy[3].

Currently, little is known about the biological correlates of MDD risk though one potential candidate is inflammation. Inflammation is part of the body's innate immune response to infection and disease and its timely response is a critical defence mechanism. However, chronic inflammation damages the body and is also known to affect the brain. Indeed, accumulating evidence suggests a possible role of inflammation in the aetiology of depression. For example, findings of cross-sectional observational studies have shown small, elevated levels of circulating inflammation biomarkers among depressed individuals[4]. Furthermore, experimental studies using animal models have shown that inducing an inflammatory state produces affective symptoms such as fatigue, social withdrawal and altered cognition[5] with similar effects also evident in humans[6]. Findings such as these suggest a possible relationship between individual risk for developing future MDD and subsequent levels of inflammatory markers, with those identified as being at higher risk for MDD in early adolescence hypothesised to have higher levels of inflammation in late adolescence than those with lower individual risk for MDD.

Study Aims

Our objective is to conduct pilot analyses (to be written-up as a 'Short Communication') to examine

associations between depression risk at age 12 and levels of inflammation at age 18 in the E-Risk cohort, using the following biomarkers: C-reactive protein (CRP), interleukin 6 (IL-6) and soluble urokinase plasminogen activator receptor (suPAR). Specific aims are:

1. To assess the relationship between MDD risk score at age 12 and measures of CRP, IL-6 and suPAR at age 18.

2. To compare age-18 levels of CRP, IL-6, and suPAR for three distinct MDD risk groups at age 12: those with MDD risk scores at extreme ends of the distribution (i.e., high- and low-risk) and those who already exhibited depressive symptoms at age 12 (and therefore were excluded from having a risk score calculation).

Notes on inflammatory markers:
Peripheral inflammation was measured from blood serum obtained at age 18. CRP is a sensitive but non-specific marker of acute inflammation that responds rapidly to underlying disease and infection[7]. IL-6 is a key cytokine involved in regulating inflammation and protecting against infection, and it is linked to chronically high inflammation due to its regulatory role[8]. suPAR is a protein that is severed from cell membranes during inflammatory conditions and provides an overall marker of immune activity and chronic inflammation[9]. The log-transformed CRP and IL-6 variables will be used as these transformed variables are more normally distributed, which has not previously been required for suPAR within the E-Risk study[9] (though we will check its distribution in the sub-set of the E-Risk sample utilised in the proposed analyses).

Statistical analyses:

Our analyses will utilise the 1,144 depression risk scores calculated at age 12 by the prediction model developed and refitted to the UK E-Risk cohort by Rocha et al[3]. The prediction model will be run using the software R and the resulting risk scores will then be exported into Stata to conduct the following analyses:

1. Using linear regression models, we will examine the association between age-12 depression risk scores and separate continuous measures of age-18 CRP, IL-6, and suPAR. Models will be adjusted for the non-independence of the twin data using the 'CLUSTER' command. Additionally, we will control for age-18 body mass index (BMI) and body temperature.

2. We will then identify three mutually exclusive groups of participants from the E-Risk sample:

    i. Those with clinically significant depressive symptoms at age 12, defined as a score of ≥20 on the Children's Depression Inventory ('depressed' group)
    ii. Those with a depression risk score ≥ the 80th percentile at age 12 ('high-risk' group)
    iii. Those with a depression risk score ≤ the 20th percentile at age 12 ('low-risk' group)

    The 20th and 80th percentile cut-offs for categorising low- and high-risk have been chosen to maximise the sample size with available inflammation data (N=150 & N=189 respectively).

3. We will check descriptive statistics for levels of CRP, IL-6, and suPAR at age 18 for each of these groups and transform variables to approximate normal distributions as required.

4. Using linear regression models, we will examine whether depression risk group membership at age 12 (depressed, high-risk, low-risk) is associated with levels of CRP, IL-6, and suPAR at age 18. Models will use the low-risk group as the reference category and will then be repeated with the depressed group as the reference. All models will be adjusted for the non-independence of twins using the 'CLUSTER' command. Additionally, we will control for age-18 BMI and body temperature.

5. As a sensitivity analysis we will repeat Steps 2-4 using the 10th and 90th percentile cut-offs for categorising participants at low and high risk for depression (N=45 & N=75 respectively).

Note: As described in Steps 1 & 4, we will adjust our linear regression models for the non-independence of twins, BMI, and body temperature. We are also considering adjusting for sex (due to sex differences in inflammation and depression prevalence) but we welcome feedback/suggestions about this given that sex is one of the predictors in the depression risk score. Note that the sample size (Step 1: N=760 & Step 2: N=393 with data available for inflammation & covariates) limits our ability to stratify by sex.

Variables Needed at Which Ages (names and labels):

Study: E-Risk Study

Age 5:

General study variables:	
FAMILYID	Unique family identifier
ATWINID	Twin A ID (ex chkdig)
BTWINID	Twin B ID (ex chkdig)
RORDERP5	Random Twin Order
RISKS	Sample Groups
COHORT	Cohort
ZYGOSITY	Zygosity

Previous depression screening for risk prediction model:	
DEPRSE5	Depression scale elder twin

Risk prediction model predictors:	
SAMPSEX	Sex of Twins: In sample
SETHNIC	Ethnicity of Twins
SESWQ35	Social class composite

Age 7:

Previous depression screening for risk prediction model:	
DEPRSE7	Depression scale elder twin

Age 10:

Previous depression screening for risk prediction model:	
DEPRSE10	Depression scale elder twin

Age 12:

Low IQ screening for risk prediction model:	
IQ12E	Pro-rated IQ (INF & MR), 12E

Previous depression screening for risk prediction model:	
CDIE12	Depression Scale - CDI – Elder

Risk prediction model predictors:	
HARME512	Child harm phase 5-12 – Elder twin
EDUCPRFE12	School performance (English & Maths average) - P12 – Elder
SISOCE12	Social isolation (categorical) - P12 – Elder
con23ec12	Have you ever run away from home and stayed away for the night? – Elder
‘Any drug use’ variable used in Rocha et al 2021. Comprised from: sub1ec12 - substance use – option 1 (upgrade of items 15, 16, and 17) – Elder sub2ec12 - substance use – option 2 (downgrade of items 03 and 07) – Elder	

'Fights' variable used in Rocha et al 2021. Comprised from: con03ec12 Do you sometimes hit someone when you are having an argument? – Elder con04ec12 Do you sometimes start fights with people? – Elder

Age 18:

Depression outcome:	
DXMDEE18	Major depressive episode, dsm4 – P18 – Elder

Inflammation:	
CRPE18_4SD	CRP concentration (mg/L) with outliers removed - P18 – Elder
lnCRP_E18_4SD	Log-transformed CRP with outliers removed – P18 – Elder
IL6_E18_4SD	Plasma IL-6 with outliers removed – P18 – Elder
lnIL6_E18_4SD	Log-transformed IL-6 with outliers removed – P18 – Elder
suPAR_E18_4SD	Plasma suPAR with outliers removed – P18 – Elder

Covariates:	
BMIE18	bmie18
BTEMPE18	Body temperature, Celsius - P18 – Elder

References cited:

1. Herrman, H., Kieling, C., McGorry, P., Horton, R., Sargent, J., & Patel, V. (2019). Reducing the global burden of depression: A Lancet–World Psychiatric Association Commission. *The Lancet, 393*(10189), e42-e43
2. Thapar, A., Collishaw, S., Pine, D. S., & Thapar, A. K. (2012). Depression in adolescence. *The Lancet, 379*(9820), 1056-67.
3. Rocha, T. B., Fisher, H. L., Caye, A., Anselmi, L., Arseneault, L., Barros, F.,… & Kieling, C. (2021). Identifying adolescents at risk for depression: A prediction score performance in cohorts based in three different countries. *Journal of the American Academy of Child & Adolescent Psychiatry, 60,* 262-273.
4. Howren, M. B., Lambkin, D. M., & Suls, J. (2009) Associations of depression with C-reactive protein, IL-1, and IL-6: a meta-analysis. *Psychosomatic Medicine, 71*, 171-186
5. Dantzer, R., O'Connor, J. C., Freund, G. G., Johnson, R. W., & Kelley, K. W. (2008). From inflammation to sickness and depression: when the inflammation system subjugates the brain. *Nature Reviews Neuroscience, 9*, 46-56.
6. Musselman, D. L., Lawson, D. H., Gumnick, J. F., Manatunga, A. K., Penna, S., Goodkin, R. S., … & Miller, A. H. (2001). Paroxetine for the prevention of depression induced by high-dose interferon alfa. *The New England Journal of Medicine, 344*, 961-966
7. Rhodes, B., Fürnrohr, B. G., & Vyse, T. J. (2011). C-reactive protein in rheumatology: biology and genetics. *Nature Reviews Rheumatology, 7,* 282-289.
8. Hunter, C.A. & Jones, S. A. (2015). IL-6 as a keystone cytokine in health and disease. *Nature Immunology, 16,* 448-457.
9. Rasmussen, L.J.H., et al., (2019). Cumulative childhood risk is associated with a new measure of chronic inflammation in adulthood. *Journal of Child Psychology and Psychiatry, 60*, 199-208

Data Security Agreement

Provisional Paper Title	Preliminary investigation of associations between individualised risk at age 12 of future depression and age-18 inflammation in a UK longitudinal cohort study
Proposing Author	Rachel Latham
Today's Date	25 February 2021

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

_RL_ 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/).

_RL_ My project has ethical approval from my institution.

_RL_ I am familiar with the EU General Data Protection Regulation (https://mrc.ukri.org/documents/pdf/gdpr-guidance-note-3-consent-in-research-and-confidentiality/), and will use the data in a manner compliant with its requirements.

_RL_ 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.

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

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

_RL_ I will not merge data from different files or sources, except where approval has been given by the PI.

_RL_ 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.

_RL_ 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.

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

_RL_ I will delete the data after the project is complete.

_RL_ **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)

____ **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: