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

Proposing Author: Alan Meehan

Author's affiliation, phone, and e-mail address:
Social, Genetic and Developmental Psychiatry (SGDP) Centre
Institute of Psychiatry, Psychology & Neuroscience,
King's College London

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

Proposed co-authors: Louise Arseneault, Daniel Stahl, Helen Fisher, Timothy Matthews, Terrie Moffitt, Avshalom Caspi, Candice Odgers

Provisional Paper Title: Individualised prediction of self-injurious thoughts and behaviours during adolescence

Date: 27th November 2019

Objective of the study and its significance:

Suicide is the second leading cause of death in adolescents and young adults (15-29 years) worldwide,[1] with rising suicide rates in recent years representing a major public health concern.[2] Suicide is the most severe endpoint of a continuum that includes various forms of self-injurious thoughts (e.g. suicidal ideation) and behaviours (e.g. deliberate self-harm, non-fatal suicide attempts).[3] These thoughts and behaviours are themselves among the strongest risk factors for later suicide attempts. For example, one UK study reported a 30-fold increase in suicide risk among individuals who had self-harmed in the previous year compared to the general population,[4] while more than a third of adolescents who report suicidal thoughts go on to make an attempt on their life.[5] At the same time, it is estimated that up to 85% of suicidal individuals make contact with health services in the 12 months leading up to their suicide attempt, often for reasons unrelated to suicidal thoughts/behaviours or mental health difficulties.[6] Consequently, there is a recognised need to identify those at risk for suicidality as early as possible to enable more targeted prevention and intervention efforts.

In an effort to inform preventive measures, previous research has successfully identified a wide array of risk factors for self-injurious thoughts and behaviours.[7,8] At the individual level, notable risk factors include: sex, with higher rates of self-harm, suicidal ideation and suicide attempts among females, but higher rates of suicide deaths among males;[9] a history of self-harm or previous suicide attempts;[4,10] psychopathology, including depressive/mood disorders, disruptive behaviour disorders, and substance use;[11-13] social isolation and loneliness;[14] impulsivity;[15] childhood abus,[8,16,17] sleep disturbances,[16] lower intelligence,[17] and several personality dimensions (e.g. high neuroticism, low extraversion).[18] Of particular note, adolescents who experience severe victimisation (e.g. physical/emotional abuse/neglect, sexual abuse, domestic violence, peer bullying) are at a markedly increased risk for self-injurious thoughts and behaviours, even after accounting for pre-existing family vulnerabilities.[19] At the family level, parental history of suicide attempts and/or psychiatric disorder, family conflict, a poor parent-child relationship, and indicators of socioeconomic adversity (e.g. household poverty, residential instability) have all been found to confer some risk of suicidality.[7,20,21] Finally, aspects of the wider community that have been linked with suicide risk include area-level socioeconomic disadvantage and a perceived lack of social support.[22,23]

Although an array of important clinical risk factors has been identified for suicidal thoughts and behaviours, meta-analytic research consistently shows that the ability of these risk factors to predict future suicidal thoughts and behaviours at an individual level remains only slightly better than chance, and well below the standard needed to justify their use in clinical decision-making.[24,25] There are three main methodological reasons that may explain the lack of robust risk prediction tools for suicide. First, the majority of initial studies examined risk factors in isolation, rather than as part of a wider risk context.[24] It is difficult to imagine that one variable could reliably distinguish individuals based on suicide risk; indeed, accurate

prediction typically involves complex combinations of risk factors across multiple domains of functioning that reflects the diverse bio-psycho-social processes underlying suicidal thoughts and behaviours. Second, most previous studies of suicide risk are cross-sectional, with retrospective reporting of both risk factors and suicide-related outcomes; thus, the extent to which these factors can predict future suicide attempts remains somewhat unclear.[26] Third, and crucially, traditional statistical techniques such as multiple logistic regression are not ideal for obtaining sufficient levels of accuracy for individual prediction. These explanatory methods essentially examine whether a given risk factor is, on *average*, associated with an increased risk for suicidality across subgroups of young people. However, there is significant individual variation among youth who are at risk of self-injurious thoughts or behaviours. For example, although the presence of depressive symptoms increases overall risk of self-harm in adolescents, not all adolescents who meet criteria for depression will deliberately harm themselves. Therefore, this risk factor, like many others, has poor predictive power when extrapolated from the group level to the individual level. Finally, most multivariate model described in the existing literature are derived using a single dataset. This introduces the possibility of 'over-fitting', whereby a model incorporates sample-specific variance to the extent that is conveys an over-optimistic estimate of the true effect in the wider population.[27] There is therefore a need to investigate the extent to which proposed predictors of suicidality can reliably classify individual risk among new or unseen cases, thus ensuring that the predictive ability obtained by a model will be consistent when tested across multiple similar-but-distinct populations.

This study aims to address these limitations by developing and internally validating a multivariable prognostic model to estimate individual risk for self-injurious thoughts and behaviours (e.g. suicidal ideation, self-harm, suicide attempts) in young people. By stratifying individual young people based on their unique risk for suicidality, these models have the potential to support and enhance clinical decision-making in at-risk youth populations. Our work will also benefit from the representativeness of the E-Risk Study, such that the resulting tools may be applicable across the wider UK population, and its prospective design, which ensures that observed risk factors preceded the onset of self-injurious thoughts and behaviours during adolescence. Specifically, beginning in the whole E-Risk sample, we will test whether a combination of well-established individual, family and community risk factors can be used to accurately distinguish those young people who go on to engage in self-injurious thoughts and behaviours from those who do not.[27] Then, given that exposure to victimisation appears to confer elevated risk for suicide-related outcomes, we will test these full-sample prediction models in a victimised sub-sample, similar to the approach used previously in E-Risk to build 'risk calculators' for psychopathology in victimised youth.[28]

Statistical analyses:

Data processing and descriptive analyses, including initial multivariable regressions, will be conducted in STATA version 15, while prediction models will be estimated using the *glmnet* package in R. All analyses will correct appropriately for familial clustering. Our analyses will comprise seven main steps:

1. To allow us to estimate a generalised prediction models for self-injurious thoughts or behaviours, we will derive a binary outcome that captures E-Risk participants who report at least one of following between ages 12 and 18 years: (i) suicidal ideation; (ii) self-harm or (iii) suicide attempts.

2. For our initial risk calculators, we will draw on more proximal information (i.e., age-12 risk factors) to identify potential predictors of self-injurious thoughts or behaviours, in order to simulate a clinical risk assessment completed at a single time-point. In preliminary analyses, bivariate regressions will be used to examine the extent to which these age-12 risk factors differentiate youth who do and do not go on to report adolescent suicidality at age 18 (i.e., group-level effects).

3. We will estimate an overall 'explanatory' model for each of our binary outcomes to assess the multivariate significance of our proposed age-12 predictor variables, as well as the relative unique contribution of each indicator to the average risk of each suicide-related outcome within E-Risk.

4. We will then test a series of prediction models for each age-18 outcome. To ensure that our models are not 'over-fitted' to the E-Risk dataset, and thus maximise their predictive ability in new or unseen cases, we will apply internal validation methods (e.g. *k*-fold cross validation, bootstrapping) to ensure that each model is 'trained' and 'tested' using independent cases or resampled data within our sample. For each model, predictive accuracy will be determined using (i) discrimination, or the

model's ability to separate victimised youth with and without psychopathology) and (ii) calibration, or the agreement between the observed and predicted outcomes.

5. Having estimated and validated these models using proximal risk factors, we will then seek to assess whether predictive accuracy can be improved by incorporating 'historical information' in our specified risk profile – that is, additional variables measured at ages 5, 7, and 10 years.

6. As E-Risk is a twin sample, it is possible that the non-independence of our observations may lead to over-optimistic prediction of suicide risk. Therefore, as a sensitivity analysis, we will re-estimate our models using ten random single-twin subsamples, before averaging the predictive performance observed across these subsamples to test whether the presence of non-independent observations has any impact on each model's predictive ability.

7. Finally, as childhood maltreatment/victimisation confers an elevated risk for self-injurious thoughts and behaviours, we will re-estimate each model in a sub-sample of victimised E-Risk youth, defined by the experience of at least one form of severe victimisation between ages 5 and 12 (n = 591), to evaluate whether predictive performance is similar for this high-risk population.

Variables Needed at Which Ages (names and labels):

Study: E-Risk Study

Age 5:

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
SAMPSEX	Sex of Twins: In sample
ZYGOSITY	Zygosity
SETHNIC	Ethnicity of Twins
SESWQ35	Social Class Composite
EXFUNCE5	Executive function – mean of Mazes, DayNight & SWM – Elder
TOMTOTE5	TOM total score – Elder
TOTEXTE5	Total Mum & Teacher Externalising Scale – Elder twin
INTISOE5	Internalising Scale – ex Soc Iso items (mum & teacher average) - P5-Elder
EMOISOE5	Emotional probs - ex Soc Iso & Somatic (mum & teacher average) - P5 -Elder
SISOE5	Social isolation scale (mum & teacher average) – P5 - Elder
ADJE5	Temperament – P5 – Elder derived by Sara Jaffee
P5ACORN	ACORN neighbourhood SES
NPROBM5	Perception of local environment
NCRIMM5	Neighbourhood personal victimisation
SOCCM5	Neighbourhood Informal Social Control
SCOHM5	Neighbourhood Social Cohesion
WARME5	Warmth towards elder twin

Age 7:

TOTEXTE7	Total Mum & Teacher Externalising Scale – Elder twin
INTISOE7	Internalising Scale – ex Soc Iso items (mum & teacher average)
EMOISOE7	Emotional probs - ex Soc Iso & Somatic (mum & teacher average) – P7 -Elder
SISOE7	Social isolation scale (mum & teacher average) – P7 – Elder
SIBWRMM7	Sibling Warmth – Derived by Sara Jaffee
ATHOME7	Atmosphere at home
FSEC7M7	Food Situation – 7 scale Recoded

Age 10:

TOTEXTE10	Total Mum & Teacher Externalising Scale – Elder twin
EMOISOE10	Internalising Scale - ex Soc Iso (mum and teacher average) - P10 - Elder

SISOE10	Social isolation scale (mum & teacher average) – P10 – Elder
LOWSC510E	Low Childhood Self-Control, 5-10, E-Twin
SIBWRM10	Sibling warmth (same as Bowes et al 2010)
ATHOME10	Atmosphere at home (same as Bowes et al., JCPP 2010)
WARME10	Warmth towards elder twin
BIOPARENTL10	Bio-parent living situation (to age 10)
FSEC7M10	Food Situation – 7 Scale Recoded
NMOVES510	N residence moves between ages 5 & 10
KIDSINFAMILY10	Number of children in the family, LHC siblings + 2

Age 12:

TOTEXTE12	Total Mum & Teacher Externalising Scale – Elder twin
EMOISOE12	Internalising Scale - ex Soc Iso (mum and teacher average) - P12 - Elder
SISOE12	Social isolation subscale (mum and teacher average) – P12 – Elder
NSOCCOHL12	Number of social parents cohabiting with bio-parent (to age 12)
IQ12E	Pro-Rated IQ (INF & MR), 12E
FCEVIDE12	Foster care or non-parental care through age 12
MONEM12	Parental Monitoring (full scale) – P12 – Elder
BFIOE12	Openness to Experience Subscale (BFI) – Elder Twin
BFICE12	Conscientiousness Subscale (BFI) – Elder Twin
BFIEE12	Extroversion Subscale (BFI) – Elder Twin
BFIAE12	Agreeableness Subscale (BFI) – Elder Twin
BFINE12	Neuroticism Subscale (BFI) – Elder Twin
ADULTEC12	Adult Involvement – Elder
HOMEM12	State of the home – Mum Intv
CHSTIM12	Child Stimulation – Mum Intv
HAPPHM12	Happy Home – Mum Intv
CHAOSM12	Chaos in the home – Mum intv
POLYVE512C	Extent of Polyvictim (Truncated @3), 5-12, E-Twin
EX_SVE12	Exposed to severe victimization (0/1), 5-12, E-Twin
ADHDANYE512	Any ADHD dx [incl meds] - P5-12 – Elder
ANYCDDX_EMT512	Any CD dx from 5 to 12, mum/tchr, Elder
MASCE12	Anxiety Scale – MASC – Elder
CDIE12	Depression Scale – CDI – Elder
SUB2EC12	Substance Use – Option 2 (Downgrade of Items 03 and 07) – Elder
SHARMSUICE12	Self-Harm/Suicidal Behaviour – P12 – Elder
PSYSYMP01E12	Psychosis Symptom Count – Verified Coding – 0, 1+ – Elder
FHSUIAM12	Anyone on list attempted suicide? – Twin's mum
FHSUICRSCR12	Fam Hist of suicide, Reeds score, Belsky 2012
FHANYPM12	Proportion of family members with valid data who have any disorder
TADHDEMT12	Total Inattentive/Hyperactive/Impulsive symptom count - (M&T average) - P12
CDTOTCRIT_EM12	Tot CD criteria met p12_Mum Elder, 2015
CDTOTCRIT_ET12	Tot CD criteria met p12_Teacher Elder, 2015
LC5M12	n change of addresses - since aged 10
	Physical health measures?

Age 18:

SUICATE18	Suicide attempted – P18 – Elder
SHARME18	Self-harm – P18 – Elder
SUIDEAE18	Any suicidal ideation 12-18 - Elder (derived by Jessie Baldwin)

References cited:

1. World Health Organization (2018). *Suicide.* https://www.who.int/mediacentre/factsheets/fs398/en/
2. Centers for Disease Control and Prevention (2017). *Web-based Injury Statistics Query and Reporting System (WISQARS)*. https://webappa.cdc.gov/sasweb/ncipc/leadcause.html
3. Bebbington, P.E., Minot, S., Cooper, C., et al. (2010). Suicidal ideation, self-harm and attempted suicide: Results from the British psychiatric morbidity survey 2000. *European Psychiatry, 25*(7), 427-431.

4. Cooper, J., Kapur, N., Webb, R., et al. (2005). Suicide after deliberate self-harm: A 4-year cohort study. *American Journal of Psychiatry, 162*(2), 297-303.
5. Nock, M.K., Green, J.G., Hwang, I., et al. (2013). Prevalence, correlates, and treatment of lifetime suicidal behaviour among adolescents. Results from the National Comorbidity Survey Replication Adolescent Supplement. *JAMA Psychiatry, 70*(3), 300-310.
6. Liu, H., Chen, L., Huang, S. (2012). Outpatient health care utilization of suicide decedents in their last year of life. *Suicide and Life-Threatening Behavior, 42*(4), 445-452.
7. Bilsen, J. (2018). Suicide and youth: Risk factors. *Frontiers in Psychiatry, 9*(540).
8. Turecki, G., & Brent, D. (2016). Suicide and suicidal behaviour. *Lancet, 387*(10024), 1227-1239.
9. Apter A, & Wasserman, D. (2006). Adolescent attempted suicide. In: R. King, & A. Apter (Eds.), *Suicide in children and adolescents* (pp. 63-85). Cambridge: Cambridge University Press.
10. Hawton, K., & van Heeringen, K. (2009). *Suicide. Lancet, 373*(9642), 1371-1381.
11. Goldston, D.B., Daniel, S.S., Erkanli, A., et al. (2009). Psychiatric diagnoses as contemporaneous risk factors for suicide attempts among adolescents and young adults: Developmental changes. *Journal of Consulting and Clinical Psychology, 77*, 281-290.
12. Meszaros, G., Horvath, L.O., & Balazs, J. (2017). Self-injury and externalizing pathology: A systematic literature review. *BMC Psychiatry, 17*(160).
13. Norstrom, T., & Rossow, I. (2016). Alcohol consumption as a risk factor for suicidal behavior: A systematic review of associations at the individual and at the population level. *Archives of Suicide Research, 20,* 489-506.
14. Calati, R., Ferrari, C., Brittner, M., et al. (2019). Suicidal thoughts and behaviors and social isolation: A narrative review of the literature. *Journal of Affective Disorders, 245*, 653-667.
15. Kasen, S., Cohen, P., & Chen, H. (2011). Developmental course of impulsivity and capability from age 10 to age 25 as related to trajectory of suicide attempt in a community cohort. *Suicide and Life-Threatening Behavior, 41*, 180-192.
16. Pigeon, W.R., Pinquart, M., & Connier, K. (2012). Meta-analysis of sleep disturbance and suicidal thoughts and behaviors. *Journal of Clinical Psychiatry, 73*(9), e1160-1167.
17. Bittár, N.H., Falkstedt, D., & Wallin, A.S. (2019). How intelligence and emotional control are related to suicidal behaviour across the life course – A register-based study with 38-year follow-up. *Psychological Medicine*. doi: 10.1017/S0033291719002423
18. Brezo, J., Paris, J., & Turecki, G. (2006). Personality traits as correlates of suicidal ideation, suicide attempts, and suicide completions: A systematic review. *Acta Psychiatrica Scandinavica, 113*(3), 180-206.
19. Baldwin, J.R., Arseneault, L., Capsi, A., et al. (2019). Adolescent victimization and self-injurious thoughts and behaviors: A genetically sensitive cohort study. *J Am Acad Child Adolesc Psychiatry*, *58*(5), 507-513.
20. Brent, D.A., Melhem, N.M., Oquendo, M., et al. (2015). Familial pathways to early-onset suicide attempt: A 5.6-year prospective study. *JAMA Psychiatry, 72*, 160-168.
21. Brent, D.A. (2016). Family risk factors for suicidal behaviour: Opportunities for early identification and intervention. In P. Courtet (Ed.), *Understanding suicide: From diagnosis to personalized treatment* (pp. 371-381). New York: Springer.
22. Cairns, J.M., Graham, E., & Bambra, C. (2017). Area-level socioeconomic disadvantage and suicidal behaviour in Europe: A systematic review. *Social Science & Medicine, 192*, 102-111.
23. Miller, A.B., Esposito-Symthers, C., & Leichtweis, R.N. (2015). Role of social support in adolescent suicidal ideation and suicide attempts. Journal of Adolescent Health, 56(3), 286-292.
24. Franklin, J.C., Ribeiro, J.D., Fox, K.R., et al. (2017). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological Bulletin, 143*(2), 187-232.
25. Kessler, R.C., Bossarte, R.M., Luedtke, A., et al. (2019). Suicide prediction models: A critical review of recent research with recommendations for the way forward. *Molecular Psychiatry*. doi:10.1038/s41380-019-0531-0
26. Mars, B., Heron, J., Klonsky, E.D., et al. (2019) Predictors of future suicide attempt among adolescents with suicidal thoughts or non-suicidal self-harm: A population-based birth cohort study. *Lancet Psychiatry, 6*(4), 327-337.
27. Steyerberg, E. (2019) *Clinical prediction models: A practical approach to development, validation, and updating*. London: Springer.
28. Meehan, A.J, Latham, R.M., Arseneault, L., et al. (2020). Developing an individualized risk calculator for psychopathology among young people victimized during childhood: A population-representative cohort study. *Journal of Affective Disorders, 262*, 90-98.

Data Security Agreement

Provisional Paper Title	Individualised prediction of self-injurious thoughts and behaviours during adolescence
Proposing Author	Alan Meehan
Today's Date	27th November 2019

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

AM   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/).

AM   My project has ethical approval from my institution.

AM   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.

AM   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.

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

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

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

AM   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.

AM   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.

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

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

____   **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 <u>never</u> 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: