GROWUPGROWOLD
HUMAN DEVELOPMENT, BIRTH TO DEATH

Moffitt & Caspi
Genes Environment Health Behavior

Concept Paper Form

Provisional Paper Title: Prevention of depression and anxiety onset in adolescence: Which at-risk groups to target?
Proposing Author: Josefien Breedvelt, PhD
Author's Email: Josefien.j.breedvelt@kcl.ac.uk
P.I. Sponsor: Prof. Andrea Danese (if the proposing author is a student or colleague of an original PI)
Today's Date: 7/27/2023

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

Objective of the Study:

Background:

Risk prediction modelling can help to identify individuals at the highest risk of developing mental health conditions, which can inform the prioritization of interventions in resource-scarce settings.[1] To date, two individual-level risk prediction models have been developed for adolescent depression prevention.[2,3] However, current individual-level risk prediction models do not indicate the size of the population with the risk factors (informing the cost of intervening) and the benefit of intervening on the incidence of mental health conditions.[4] As such, current models have limited utility to plan preventative interventions. An alternative approach is to select risk factors that identify the smallest possible subgroup of children and adolescents at greatest risk for depression or anxiety onset.[4,5,6]

In addition, current prediction models are focused and limited to depression. However, there is high comorbidity between depression and anxiety disorders, which increases during adolescence.[7] The same risk factors (i.e., abuse, family history of psychiatric disorder, history of physical disease) predict risk of both depression and anxiety disorders.[8] In focus groups with service users, we found that they preferred supporting the development of transdiagnostic models for depression and anxiety disorders rather than two separate models for each condition. As such, future prediction modeling studies should jointly model the risk for depression and anxiety disorders in adolescence.

Data analysis methods:

Is it possible to identify the smallest group of children at the highest risk of developing depression or anxiety in order to inform future development of targeted depression and anxiety prevention approaches in children and adolescents?

1. **Selecting predictors**

We start by identifying a smaller but comprehensive set of predictors of depression or anxiety based on the longer list of putative predictors included in the variable request form.
We will use the (automated) back-step procedure at $p<0.10$ in a multivariable Poisson model to regress the binary outcome variable of depression or anxiety (generalized anxiety disorder or

PTSD) onset at age 18 year on the complete set of risk factors (predictor variables).

This will select a smaller set of predictor variables where each predictor has a unique and significant contribution to predicting depression and anxiety.

As we are interested in primary prevention, which is the prevention of first onset depression and/or anxiety, we will first analyze predictors in a sample without a known past experience of high depressive or anxiety symptoms. Known past experience of high depressive or anxiety symptoms is defined by experiencing above cut-off symptoms at ages 5, 7 or 10 as measured with the depressive symptom/anxiety symptom questionnaire from the CBCL (using the 93th percentile as cut-off) and above cut-off symptoms at age 12 as measured with the CDI and MASC.

The cut-off criteria are further detailed in the variable request form and are based on previous work with E-RISK using these cut-offs. For instance, the 93th percentile for the CBCL was used by Rocha et al., 2021[3] to determine previous high depressive symptomatology.

If our sample size is insufficient, i.e., we have limited power to run our analysis with the aforementioned subsample, an analysis on the whole sample (including those with a prior history of high symptoms) will be performed. Limited power is defined as having an event per variable ratio at 10 or lower.[9] The event per variable ratio is calculated by dividing the number of events (new cases of onset) by the number of risk factors considered in developing the prediction model (as included in our variable request form). Should we proceed with the aforementioned analysis due to low power, we will include a classifier for lifetime experience of anxiety or depression to account for past experience of high depressive and/or anxiety symptoms.

2. **Calculating the Exposure Rate (ER), Incidence Rate Ratio (IRR), Population Attributable Fraction (PAF) and Number Needed to Treat (NNT)**

The ER, IRR, PAF and NNT will first be calculated separately for all predictor variables that significantly predict depression and anxiety. These terms, and how they are calculated are further described below.

The formulae below include the following groups, where:
Group A = total number of participants exposed to risk factor
Group B = total number of participants not exposed to risk factor
Group C = total number of new cases of onset in group exposed to risk factor
Group D = total number of new cases of onset in group not exposed to risk factor
Group E = total sample

Exposure rate (ER) = Group A / Group E

The ER is the percentage of children and young people with the risk factor. The ER is calculated by dividing the number of participants exposed to the risk factor by the total sample.

Incidence Rate Ratio (IRR) = ((Group C / Group E) / (Group D / Group E))

The IRR equates to the incidence rate of participants exposed to the risk factor divided by incidence rate of participants not exposed to the risk factor. The incidence rate is the number of new cases of onset (i.e. incidence) divided by the total number of participants at risk in the sample. An incidence rate ratio higher than 1 suggests a higher level of risk in the group exposed to the risk factor, and an incidence rate ratio lower than 1 indicates a risk reduction in the exposed group. In order to calculate the IRR, a univariate Poisson regression will be conducted where the outcome (onset or non-onset status) is regressed on the risk factor.

Population Attributable Fraction (PAF) = (ER x (RR - 1)) / (1 + ER x (RR - 1))

Where:
ER = Exposure Rate (exposed proportion of the population)
RR = Risk Ratio; the risk of onset in the exposed group divided by the risk of onset in the unexposed group = ((Group C / Group A) / (Group D / Group B))

The PAF is the percentage of cases of onset that can be attributable to exposure to a risk factor (or combination thereof).[10] It indicates by how many percentage points the incidence of anxiety or depression would be reduced if the effect of the predictor is cancelled out. In other words, the PAF provides the largest estimate of the impact of a preventative intervention for each predictor, assuming causality.

The Absolute Risk Reduction (ARR) = ((Group C / Group A) - (Group D / Group B))

The absolute risk reduction is the risk of the outcome in the exposed group minus the risk of the outcome in the unexposed group. The absolute risk reduction will be obtained by regressing the incidence on the risk factor. The ARR will be calculated in order to compute the number needed to treat.

Number needed to treat (NNT) = 1/ ARR

The NNT is calculated by the inverse of the ARR. The NNT is number of people who would need treatment before one case of depression or anxiety is prevented, assuming causality.

3. Optimising configuration of risk factors

We assume that combining multiple risk factors allows us to identify a group that best satisfies the optimal values of IRR (small), PAF (large), NNT (small), ER (small). For example, by combining two risk factors (bullying and maltreatment) the exposure rate will be smaller (as fewer children will have experienced both) but the PAF may be larger (as experiencing both factors may heighten the risk of onset).

In order to identify which and how many risk factors can be best combined, we propose a stepwise process as described below:

1. We will create a table displaying all risk factors significant at $p<0.10$ (as per the Poisson regression) and their associated IRR, PAF, NTT and ER.
2. Then, one risk factor will be identified that best satisfies the optimal values of IRR (small), PAF (large), NNT (small), ER (small).
3. The aforementioned risk factor is modelled in a Poisson regression with each additional risk factor that was significant at $p<0.10$.
4. The IRR, PAF, NNT and ER are then calculated for the combination of the first and additional risk factor.
5. The IRR, PAF, NNT and ER for the combination of two risk factors will be presented in a table and the optimal combination will be identified that best satisfies the values of IRR (small), PAF (large), NNT (small), ER (small).
6. If an optimal combination of two risk factors is identified, a third risk factor may be added. Here, we will model the not yet included risk factors together with the two already included risk factors in a multivariable Poisson regression and assess if a combination of risk factors with optimal values of IRR (small), PAF (large), NNT (small), ER (small) can be identified.

The results will be displayed in tables and a decision tree similar to a multi-criterion Classification and Regression Tree (CART) analysis diagram (see Figure 1).[11–13] This will aid the visual

interpretation of how different combinations of predictor variables can affect the IRR, PAF, NNT and ER.

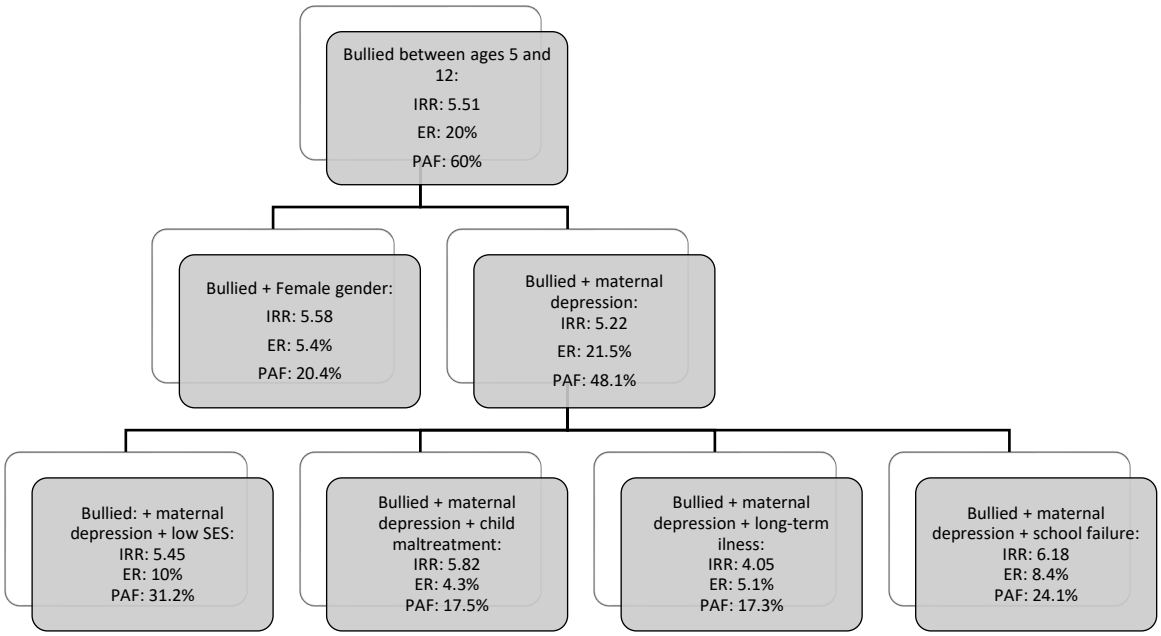

Figure 1. Adapted from Figure 1. Smit et al., 2007[12] – "selecting combinations of risk indicators where the incidence rate ratio is greater than 5 and the exposure rate is lower than 10% while maintaining a population attributable fraction as high as possible"

4. Number of models

We will first create one transdiagnostic model to identify risk factors for being diagnosed with depression, anxiety or both at age 18.

For our sensitivity analyses, we will evaluate whether predictors vary by anxiety and depression, and if separating models by anxiety or depression will improve performance on IRR, ER, PAF and NNT. We will also evaluate whether our models vary if we include PTSD (in the anxiety disorder model) or not.

If we have enough statistical power, we will also create a separate model for PTSD at age 18, which includes a population limiting factor of having experienced trauma in childhood (prior to onset).

The primary target is to identify a subgroup that may potentially be best targeted with a preventative approach. This subgroup may be targeted with interventions that target modifiable risk factors. However, not all risk factors in the model may be modifiable.

Significance of the Study (for theory, research methods or clinical practice):

Current individualised prediction modelling techniques don't indicate the extent of the population that is affected by a risk factor (e.g. 'at risk') and the mental health benefit of intervening in a high-risk group.[14] This lack of "cost-benefit" information is a known barrier to implementing current risk modelling in clinical practice.[15]

This research proposes a complementary approach to current risk prediction modelling by adding in population-level cost-benefit indicators. A population-level approach will be taken that has been

largely applied in medicine with limited uptake in preventive psychiatry and youth prevention.[4,6,11]

By creating an improved understanding of whom to target with a preventative intervention, this study may aid policymakers and commissioners in identifying potential target groups for intervention. This may, in future, lead to more targeted delivery of preventative interventions. The selected risk factors may either be directly targeted/modified by specific interventions or provide means to identify at-risk subgroups in need of more generic early interventions.

To facilitate further impact of our research, we aim to create an RShiny web-app to display our results and allow straightforward interpretation. Prior to further development of the app, we will conduct extensive scoping and user testing with our policy and practice advisory group (consisting of service providers, public health officials and third-sector organisations involved in the provision of preventative services in the UK). The web-app will run on and only display published model estimates.

Variables needed at which ages:

The included variables represent the variables that were previously studied in risk prediction models and/or are supported by recent (systematic) reviews.[2,3,16–20]

Variable	Label	Rationale for request and references where variable has been included in prior reviews and/or prediction models
Age 5		
FAMILYID	Unique family identifier	
ATWINID	Twin A ID (ex chkdig/ Primary Elder Twin ID)	
LILLEL5*~	Any long-term illness - health records	We have requested LILLEL5 and NOILLEL7 as we would like to create a lifetime variable of having had a long-term illness in the model. Experience of illness and disability has been included in prior prediction models of depression.[16,17]
SESWQ35*	Social Class Composite	If possible, we would like to have the three variables that comprise SES at age 5 separately. All these variables were previously separately included in prediction models and reviews of prognostic factors (Total household income:[2,21–23], Education (Mother)[16,17], Employment [16,17,19]). Having the variables separately could improve the prediction model and it is also simpler to implement and communicate.
SAMPSEX*	Sex of Twins: In sample	[2,3,16–19]
Age 7		
A7STATUS	Interview outcome	
NOILLEL7*~	No Illness P5 to P7 - Elder	We have requested LILLEL5 and NOILLEL7 as we would like to create a lifetime variable of having had a long-term illness in the model. Experience of illness and disability has been included in prior prediction models.[16,17]
TRF3E7*~	Ever been referred for a special education programme or social services? - Elder twin	Special educational needs is included as a risk factor for depression and anxiety. We have requested TRF3E7 and SE10M12 to develop a longitudinal variable of ever having been identified as having special educational needs.[24]
Age 12		
A12STATUS	Interview outcome	

SchlFailE12*	School failure (No). Evaluation of sample's distribution of English/Math performance at age 12, considering those below the 20th percentile as "failing at school"=1; otherwise=0	3
PH12URB_CAT3*	Urbanicity (3 cat) - P12	25
con23ec12*	Responses to the dichotomous question: "Have you run away from home and stayed away for the night?"; positive answers were classified as "1", and negative as "0"	3
BULLSEVE12*	Bullying victim to Age 12 - Elder	15
SocIso01E12*	Social isolation (0). Combination of CBCL and TRF items on social isolation was pooled into a 3 strata categorical social isolation variable (low, moderate and high social isolation), and then reclassified into a dichotomized variable: high social isolation=1 and low/moderate=0	3,16
FightsDicE12*	Fights. Combination ("AND" rule) of two dichotomous questions: "Do you sometimes hit someone when you are having an argument?" and "Do you sometimes start fights with people?"; positive answers to both questions were classified as "1"; otherwise, "0"	3
harme512*	Childhood maltreatment (None). Prospectively obtained variable for sexual/physical abuse up to age 12 based on mother reports, researcher observations, and social services referral information, coded as none, probable, or definite; inserted as a categorical variable into the model	2,3,16
LEVNTM12*	No. Stressful Life Events - Since Twins Aged 10	Stressful life events included in prior models as predictor variable. In Stephens et al., 2023 a 4-item stressful life event screener assessing the categories "death of a family member" and "parental separation or discord".[2]
RecDepM512*	Mum recurrent depression, 2 or more 5 to 12	2,17
DXGADM12*	Gen Anxiety Disorder - DSM4 - Mother	2,17
SE10M12*~	Special education service or SENCO - Elder	Indicator of SEN / disability (see comment above re: longitudinal SEN status variable).[24]
anysubusee12*	Drug use (No). Dichotomous variable combining responses to dichotomous questions about any lifetime use of alcohol, tobacco, cannabis, pills and inhalants; any positive answer="1"; otherwise="0"	3
DepScreenE512"	As defined by Rocha et al., 2021.[3] *"Any evidence ("OR" rule) of depressive symptoms, assessed at ages 5, 7 and 10 by a depression subscale derived from a combination of mother and teacher CBCL for emotional problems, using the 93th percentile as cut point, and at age 12 by self-reported CDI scores (with a clinical cut-off >= 20)."*	2,3,16,17 This screen, as also used by Rocha et al., 2021 will be used to determine whether the sample did or did not screen positive for previous high depressive symptomatology.
CDIE12*	Depression Scale - CDI - Elder	Depressive symptoms are included in this variable request form as sub-threshold variation in symptoms may be present and affect risk of developing depression and/or anxiety.
MASCE12*	Anxiety Scale - MASC - Elder	2,16,17
		We would also like to derive a variable similar to DepScreenE512 for experiencing high levels of anxiety symptoms at ages 5, 7 and 10 and scoring above cut-off on the MASC at age 12
Age 18		

DXMDEE18^	Major depressive episode, dsm4 - P18 - Elder	
DXGADE18^	Gen Anxiety Disorder, dsm4_based - P18 - Elder	
DXPTSD5CUE18^	PTSD Current dx, dsm5 - P18 - Elder	

*= predictor
^= outcome variable
" = sample specifier (to determine whether sample did or did not have depression and/or anxiety in past)
~= will be combined with other variable to create a longitudinal variable

References cited:

1. Meehan AJ, Lewis SJ, Fazel S, et al. Clinical prediction models in psychiatry: a systematic review of two decades of progress and challenges. *Mol Psychiatry*. 2022;27(6):2700-2708. doi:10.1038/s41380-022-01528-4

2. Stephens A, Allardyce J, Weavers B, et al. Developing and validating a prediction model of adolescent major depressive disorder in the offspring of depressed parents. *J Child Psychol Psychiatry*. 2023;64(3):367-375. doi:https://doi.org/10.1111/jcpp.13704

3. Rocha TBM, Fisher HL, Caye A, et al. Identifying Adolescents at Risk for Depression: A Prediction Score Performance in Cohorts Based in 3 Different Continents. *J Am Acad Child Adolesc Psychiatry*. 2021;60(2):262-273. doi:10.1016/j.jaac.2019.12.004

4. Cuijpers P, Smit F, Furukawa TA. Most at-risk individuals will not develop a mental disorder: the limited predictive strength of risk factors. *World Psychiatry Off J World Psychiatr Assoc WPA*. 2021;20(2):224-225. doi:10.1002/wps.20852

5. Schoevers RA, Smit F, Deeg DJH, et al. Prevention of late-life depression in primary care: Do we know where to begin? *Am J Psychiatry*. 2006;163(9):1611-1621. doi:10.1176/appi.ajp.163.9.1611

6. Smit F, Ederveen A, Cuijpers P, Deeg D, Beekman A. Opportunities for Cost-effective Prevention of Late-Life Depression: An Epidemiological Approach. *Arch Gen Psychiatry*. 2006;63(3):290-296. doi:10.1001/archpsyc.63.3.290

7. Merikangas KR, He JP, Brody D, Fisher PW, Bourdon K, Koretz DS. Prevalence and treatment of mental disorders among US children in the 2001-2004 NHANES. *Pediatrics*. 2010;125(1):75-81. doi:10.1542/peds.2008-2598

8. Arango C, Dragioti E, Solmi M, et al. Risk and protective factors for mental disorders beyond genetics: an evidence-based atlas. *World Psychiatry*. 2021;20(3):417-436. doi:10.1002/wps.20894

9. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol*. 1996;49(12):1373-1379. doi:10.1016/s0895-4356(96)00236-3

10. Rothman KJ. *Epidemiology : An Introduction* . 2nd editio. Oxford University Press; 2012.

11. Monshouwer K, Smit F, Ruiter M, et al. Identifying target groups for the prevention of depression in early adolescence: the TRAILS study. *J Affect Disord*. 2012;138(3):287-294. doi:10.1016/j.jad.2012.01.026

12. Smit F, Comijs H, Schoevers R, Cuijpers P, Deeg D, Beekman A. Target groups for the prevention of late-life anxiety. *Br J Psychiatry J Ment Sci*. 2007;190:428-434. doi:10.1192/bjp.bp.106.023127

13. Schoevers RA, Smit F, Deeg DJH, et al. Prevention of late-life depression in primary care: do we know where to begin? *Am J Psychiatry*. 2006;163(9):1611-1621. doi:10.1176/ajp.2006.163.9.1611

14. Cuijpers P, Smit F, Furukawa T. Most at-risk individuals will not develop a mental disorder : the limited predictive strength of risk factors. 2021;(June):224-225. doi:10.1002/wps.20852

15. Baldwin H, Loebel-Davidsohn L, Oliver D, et al. Real-World Implementation of Precision Psychiatry: A Systematic Review of Barriers and Facilitators. *Brain Sci*. 2022;12(7). doi:10.3390/brainsci12070934

16. Bellón JÁ, de Dios Luna J, King M, et al. Predicting the onset of major depression in primary care: international validation of a risk prediction algorithm from Spain. *Psychol Med*. 2011;41(10):2075-2088. doi:10.1017/S0033291711000468

17. King M, Walker C, Levy G, et al. Development and validation of an international risk prediction algorithm for episodes of major depression in general practice attendees: the PredictD study. *Arch Gen Psychiatry*. 2008;65(12):1368-1376. doi:10.1001/archpsyc.65.12.1368

18. Bellón JÁ, de Dios Luna J, King M, et al. Predicting the onset of major depression in primary care: international validation of a risk prediction algorithm from Spain. *Psychol Med*. 2011;41(10):2075-2088. doi:10.1017/S0033291711000468

19. King M, Bottomley C, Bellón-Saameño JA, et al. An international risk prediction algorithm for the onset of generalized anxiety and panic syndromes in general practice attendees: predictA. *Psychol Med*. 2011;41(8):1625-1639. doi:10.1017/S0033291710002400

20. Fusar-Poli P, Correll CU, Arango C, Berk M, Patel V, Ioannidis JPA. Preventive psychiatry: a blueprint for improving the mental health of young people. *World Psychiatry*. 2021;20(2):200-221. doi:https://doi.org/10.1002/wps.20869

21. Allen J, Balfour R, Bell R, Marmot M. Social determinants of mental health. *Int Rev Psychiatry Abingdon Engl*. 2014;26(4):392-407. doi:10.3109/09540261.2014.928270

22. Patel V, Burns JK, Dhingra M, Tarver L, Kohrt BA, Lund C. Income inequality and depression: a systematic review and meta-analysis of the association and a scoping review of mechanisms. *World Psychiatry*. 2018;17(1):76-89. doi:10.1002/wps.20492

23. Ridley M, Rao G, Schilbach F, Patel V. Poverty, depression, and anxiety: Causal evidence and mechanisms. *Science*. 2020;370(6522):eaay0214. doi:10.1126/science.aay0214

24. Totsika V, Liew A, Absoud M, Adnams C, Emerson E. Mental health problems in children with intellectual disability. *Lancet Child Adolesc Health*. 2022;6(6):432-444. doi:10.1016/S2352-4642(22)00067-0

25. van der Wal JM, van Borkulo CD, Deserno MK, et al. Advancing urban mental health research: from complexity science to actionable targets for intervention. *Lancet Psychiatry*. 2021;8(11):991-1000. doi:10.1016/S2215-0366(21)00047-X