

E-Risk Study Concept Paper Form

Response was completed on 23-07-2025 21:19.

Record ID

51

1. Collaborating researchers

Please note:

Once approved, a formal data use agreement will be required between King's College London and the university or research organisation that employs any collaborator having access to the data if they are not a member of staff, a student or affiliate of King's College London. This needs to be signed by both universities/organisations before data access can be granted.

For projects carried out by a student (e.g., MSc/MA, MPhil/PhD, clinical doctorate), the lead applicant should be the student's supervisor at the same university, and the student should be named as the student collaborator requiring access to the data.

If you have additional collaborators, please name them below and indicate whether they need to have access to the data. It would be common, for instance, for other researchers to see summary results of analyses and act as co-authors on your paper without having access to the data. You will not be permitted to share the dataset except with those indicated in the table as requiring access.

| Applicable? | Category | Name | Email address | University/organisation | Needs access to data for analysis? |
|---|---|--------------|--|-------------------------|--|
| | Applicant (lead researcher) | Xinyang Yu | xinyang.1.yu@kcl.ac.uk | King's College London | <input checked="" type="radio"/> Yes <input type="radio"/> No |
| <input type="radio"/> Applicable <input checked="" type="radio"/> Not applicable | Student collaborator (if data is for their dissertation/thesis) | | | | |
| <input checked="" type="radio"/> Applicable <input type="radio"/> Not applicable | E-Risk Sponsor (if applicant is not an E-Risk investigator) | Helen Fisher | helen.2.fisher@kcl.ac.uk | King's College London | <input type="radio"/> Yes <input checked="" type="radio"/> No |
| Are there additional collaborators to add? | | | <input checked="" type="radio"/> Yes <input type="radio"/> No | | |
| If yes, how many additional collaborators would you like to add? | | | <div>9</div> <div>▼</div> | | |

| Category | Name | Email address | University/organisation | Needs access to data for analysis? |
|-----------------------|---------------------|-------------------------------|-------------------------|--|
| Other collaborator #1 | Sylvane Desrivieres | sylvane.desrivieres@kcl.ac.uk | King's College London | <input checked="" type="radio"/> Yes <input type="radio"/> No |
| Other collaborator #2 | Louise Arseneault | louise.arseneault@kcl.ac.uk | King's College London | <input type="radio"/> Yes <input checked="" type="radio"/> No |
| Other collaborator #3 | Terrie Moffitt | terrie.moffitt@duke.edu | Duke University | <input type="radio"/> Yes <input checked="" type="radio"/> No |
| Other collaborator #4 | Avshalom Caspi | avshalom.caspi@duke.edu | Duke University | <input type="radio"/> Yes <input checked="" type="radio"/> No |
| Other collaborator #5 | Chloe Wong | chloe.wong@kcl.ac.uk | King's College London | <input type="radio"/> Yes <input checked="" type="radio"/> No |
| Other collaborator #6 | Jonathan Mill | j.mill@exeter.ac.uk | University of Exeter | <input type="radio"/> Yes <input checked="" type="radio"/> No |
| Other collaborator #7 | Karen Sugden | karen.sugden@duke.edu | Duke | <input type="radio"/> Yes <input checked="" type="radio"/> No |
| Other collaborator #8 | Benjamin Williams | benjamin.s.williams@duke.edu | Duke | <input type="radio"/> Yes <input checked="" type="radio"/> No |

| | | | | |
|--------------------------|--------------|-----------------------------|-------------------------|--|
| Other collaborator #9 | Eilis Hannon | E.J.Hannon@exe ter.ac.uk | University of Exeter | <input type="radio"/> Yes <input checked="" type="radio"/> No |
| | | | | |

Applicants: If you would like to continue your application later, please press the "Save and return later" button below. Please copy or write down the Return code provided.

To return later, you may click on "Returning?" on the top right of the screen in the E-Risk Concept Paper Form link, which is the same link that was used to access this form: <https://redcap.link/ERiskConceptPaperForm>

2. The project proposal

Note: Please provide sufficient detail to enable the committee to review your proposal. Please be as specific as possible about the project aims and analysis methods as once approved this concept paper will be posted publicly and thus will act as a form of pre-registration of your project. Expand boxes as required.

| | |
|--|---|
| Title of project | The Epigenetic Architecture of Brain and Behaviour: Normative Modelling from ENIGMA-Epigenetics |
| Background and rationale for project (approx. 300 - 1000 words) | <p>Mental disorders are among the leading global contributors to years lived with disability, yet our ability to detect early signs of vulnerability and intervene before clinical onset remains limited. To move beyond symptom-based classification and improve early risk stratification, we need biologically grounded, developmentally informed biomarkers that reflect individual differences in brain and behavioural trajectories across the lifespan.</p> <p>DNA methylation (DNAm) is a key epigenetic mechanism that mediates the influence of genetic and environmental factors on gene expression (Moore et al., 2013). Altered DNAm profiles have been associated with a wide range of psychiatric, neurodevelopmental, and neurodegenerative disorders (Grezenko et al., 2023; Reichard & Zimmer-Bensch, 2021). While epigenetic clocks have offered insights into biological ageing (Hannum et al., 2013; Horvath, 2013), they typically focus on limited CpG subsets, adult populations, and aggregate measures, lacking spatial and developmental resolution.</p> <p>In contrast, normative modelling offers a powerful statistical framework to characterise inter-individual variation by defining age-specific reference trajectories and quantifying deviations from them (Marquand et al., 2016; Rutherford et al., 2023). Originally applied in neuroimaging to construct lifespan brain charts (Bethlehem et al., 2022; Ge et al., 2024; L. Sun et al., 2025), this approach has yet to be systematically applied to genome-wide epigenetic data across the human lifespan. Developing normative epigenetic models could transform our ability to detect atypical biological development before clinical symptoms emerge, with clear implications for early intervention and personalised care.</p> <p>Building on Our Contributions to Epigenetics and Brain Phenotypes: The proposed project builds directly on our prior and ongoing work within the IMAGEN and ENIGMA-Epigenetics initiatives. As PI of this application, I (SD) lead the ENIGMA-Epigenetics Working Group, a global collaboration of 20+ cohorts and >8,000 individuals aged 12 to 87 years, which aims to identify blood-based epigenetic biomarkers of brain structure and psychiatric risk. I am also UK-PI for the IMAGEN study, a longitudinal cohort tracking adolescent brain and behavioural development.</p> <p>Our research has produced several landmark findings at the intersection of genetics, epigenetics, and brain imaging:</p> <ul style="list-style-type: none"> • We were the first to report that blood DNAm at a specific gene locus predicts both brain |

| | |
|---|--|
| | <p>activation and future alcohol misuse in adolescents (Ruggeri et al., 2015), a study publicly highlighted by the American Psychiatric Association.</p> <ul style="list-style-type: none"> • In collaboration with ENIGMA-Epigenetics, we identified genome-wide DNAm correlates of hippocampal volume, implicating genes involved in memory, metabolism, and stem cell maintenance (Jia et al., 2021). • Our longitudinal EWAS within IMAGEN revealed how DNAm mediates the impact of negative life events on adolescent brain structure and ADHD-related behaviours (Y. Sun et al., 2022). • Most recently, our meta-analysis of DNAm and cortical morphology in 7,400 individuals across 20 cohorts contributing to ENIGMA-Epigenetics identified distinct CpGs and differentially methylated regions (DMRs) associated with cortical thickness (CT) and surface area (SA). These loci showed minimal overlap, consistent with divergent epigenetic regulation, and were enriched in pathways linked to neural signalling, metabolism, and psychiatric traits (manuscript in preparation). Mendelian randomisation analyses suggest potential causal effects of DNAm on cortical morphology and psychiatric disorders. <p>A New Paradigm: Normative Epigenetic Trajectories Across the Lifespan: Despite these advances, we currently lack a normative framework to model age-related changes in DNAm at the genome-wide level, an essential step to detect when and how individuals diverge from typical developmental patterns.</p> |
| Project aims / objectives | <p>This project aims to fill that gap by developing normative epigenetic models across the human lifespan, using harmonised data from IMAGEN, ENIGMA-Epigenetics cohorts, and other cohorts, including the E-Risk cohort.</p> <p>The E-Risk cohort (age 18) will contribute to the training sample for normative models, enhancing resolution during late adolescence, and will also serve for exploratory phenotype-DNAm analyses.</p> |
| Brief statement of your hypothesis | <p>Aim 1: Normative DNAm trajectories across the lifespan We hypothesise that DNA methylation at individual CpG sites follows developmentally dynamic, non-linear, and biologically meaningful trajectories, with potential sex-specific patterns. By integrating data from multiple international cohorts, we will construct normative reference models that capture these trajectories. The inclusion of the E-Risk cohort at age 18 will provide critical resolution during late adolescence, a period sometimes underrepresented in lifespan studies.</p> <p>Aim 2: Clusters of CpGs and associations with health and behaviour We further hypothesise that groups of CpGs sharing similar age-related patterns index coordinated biological processes relevant to health and behaviour. Specifically, we expect that trajectory-informed CpG clusters will be associated with substance use, psychiatric symptoms, and health-related outcomes such as body mass index. Within the E-Risk sample, we will test whether deviations from normative DNAm trajectories are linked to these phenotypic domains.</p> |
| Data analysis methods to be used <i>(approx. 100 - 500 words)</i> | <p>DNAm data from the E-Risk cohort will be used in two complementary ways within this project: (1) to support normative modelling of age-related DNAm trajectories across the lifespan, and (2) to conduct exploratory analyses of health and environmental associations based on trajectory-derived features.</p> <p>1. Normative Modelling To address our first hypothesis, we will model age-related DNA methylation trajectories at the single-CpG level across the lifespan. Trajectories will be estimated using Generalised Additive Models for Location, Scale, and Shape (GAMLSS; Bethlehem et al., 2022), which allow flexible modelling of both the mean and variance of methylation levels and can capture non-linear developmental patterns. E-Risk DNAm data (age 18) will be harmonised with data from other population-based cohorts covering a broad age range, and included in the training sample.</p> <p>Age will be the primary predictor, with separate models for males and females to identify</p> |

sex-specific trajectories, and pooled models including age-by-sex interactions to formally test differences. All models will account for estimated cell-type composition, technical covariates, including batch and array type, and cohort as a random effect, to control for inter-cohort differences.

To ensure comparability across different DNA methylation platforms and maintain biological signal, we will implement a multi-step harmonisation procedure. Probes with known cross-reactivity, SNP overlap, or poor reproducibility will be excluded. Probe reliability will be assessed using intra-class correlation coefficients (ICCs) for replicate samples, retaining only probes demonstrating convincing reliability. We will also incorporate validated probes sets from prior reliability studies (Sugden et al., 2020).

Preprocessing will follow standardised pipelines with appropriate normalisation applied across cohorts. Residual batch and platform effects will be adjusted using methods such as ComBat, with study and platform included as covariates or random effects in all statistical models to account for residual variation.

To validate the biological signal and confirm that harmonisation preserves true signal, we will test established positive controls (e.g., age-associated CpGs such as sites in the ELOVL2 promoter (Ochana et al., 2025) across platforms. As sensitivity analyses, we will also evaluate stability of results within cohorts measured longitudinally on the same platform, confirming that harmonisation does not attenuate biological effects.

Model selection will use penalised splines and information criteria, and model reliability will be evaluated through cross-validation, jack-knife resampling, and leave-cohort-out analyses. Multiple testing correction at the CpG level will ensure appropriate control of false positives.

Given the focus on normative modelling, sample size considerations relate to the precision of trajectory estimates rather than classical hypothesis testing. Large reference cohorts improve the reliability of predicted methylation values and the sensitivity to detect meaningful deviations (Rutherford et al., 2022). Across ENIGMA-Epigenetics and other collaborating cohorts, we anticipate >20k individuals spanning ages 0-80 years. Pilot analyses in ~4,000 participants have already revealed distinct DNAm trajectory patterns (see Figure 1).

**** insert Figure 1 here ****

Figure 1: DNAm trajectory patterns across age in pilot cohort (approximately N = 4,000).

The inclusion of the E-Risk cohort at age 18 will further enhance resolution in this developmental window. We will also consider pooling narrow age ranges, when necessary, to stabilise estimates. Model robustness will be assessed through cross-validation and bootstrapping, ensuring reliable normative estimates even in narrower age bands.

2. Exploratory Analysis of Environmental and Health Associations in E-Risk

For the second hypothesis, we will characterise groups of CpGs with similar developmental trajectories. Trajectory features, including intercept, slope, curvature, and inflection points, will be extracted for each CpG. These features will then be subjected to unsupervised clustering methods, such as k-means, hierarchical clustering, or Gaussian mixture models. The number of clusters will be determined using internal validity indices and bootstrapping to ensure stable solutions.

For each cluster, we will calculate a summary score, either as the mean methylation across CpGs or the first principal component. Associations with health and behavioural outcomes will be tested using appropriate models:

- Continuous outcomes, such as body mass index or substance use (e.g., smoking, alcohol, marijuana, drug dependence), will be examined using linear mixed-effects models.
- Binary outcomes, such as, psychiatric diagnoses or environmental exposures (e.g.,

| | |
|--|---|
| | <p>socioeconomic disadvantage), will be examined using logistic regression.</p> <p>All models will include sex, ancestry, cell-type composition, and technical covariates as fixed effects, with cohort and, for twin data, family structure included as random effects.</p> |
| Significance for theory, research methods, or clinical practice | <p>From a theoretical perspective, using E-Risk data will improve the modelling of DNAm patterns during adolescence - a key developmental period that is often underrepresented in lifespan studies. By providing DNAm and detailed phenotypic data at age 18, the cohort will help clarify how epigenetic patterns during this stage might be shaped by neurodevelopmental processes and environmental exposures. Exploratory analyses will examine whether groups of CpG sites, identified based on their age-related patterns across cohorts, are associated with psychiatric symptoms and substance use within the E-Risk sample.</p> <p>From a methodological perspective, the project will develop a framework for normative modelling of genome-wide DNAm data across the lifespan. Including E-Risk in the training sample will improve the accuracy of trajectory estimates during late adolescence and increase the developmental coverage of the reference models. These reference models will then be applied in external datasets to investigate their potential clinical relevance.</p> |
| References cited | <p>Bethlehem, R. a. I., Seidlitz, J., White, S. R., Vogel, J. W., Anderson, K. M., Adamson, C., Adler, S., Alexopoulos, G. S., Anagnostou, E., Areces-Gonzalez, A., Astle, D. E., Auyeung, B., Ayub, M., Bae, J., Ball, G., Baron-Cohen, S., Beare, R., Bedford, S. A., Benegal, V., ... Alexander-Bloch, A. F. (2022). Brain charts for the human lifespan. <i>Nature</i>, 604(7906), Article 7906. https://doi.org/10.1038/s41586-022-04554-y</p> <p>Ge, R., Yu, Y., Qi, Y. X., Fan, Y., Chen, S., Gao, C., Haas, S. S., New, F., Boomsma, D. I., Brodaty, H., Brouwer, R. M., Buckner, R., Caseras, X., Crivello, F., Crone, E. A., Erk, S., Fisher, S. E., Franke, B., Glahn, D. C., ... Yu, K. (2024). Normative modelling of brain morphometry across the lifespan with CentileBrain: Algorithm benchmarking and model optimisation. <i>The Lancet Digital Health</i>, 6(3), e211-e221. https://doi.org/10.1016/S2589-7500(23)00250-9</p> <p>Grezenko, H., Ekhatov, C., Nwabugwu, N. U., Ganga, H., Affaf, M., Abdelaziz, A. M., Rehman, A., Shehryar, A., Abbasi, F. A., Bellegarde, S. B., Khaliq, A. S., Grezenko, H., Ekhatov, C., Nwabugwu, N. U., Ganga, H., Affaf, M., Abdelaziz, A. M., Rehman, A., Shehryar, A., ... Khaliq, A. S. (2023). Epigenetics in Neurological and Psychiatric Disorders: A Comprehensive Review of Current Understanding and Future Perspectives. <i>Cureus</i>, 15(8). https://doi.org/10.7759/cureus.43960</p> <p>Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., Klotzle, B., Bibikova, M., Fan, J.-B., Gao, Y., Deconde, R., Chen, M., Rajapakse, I., Friend, S., Ideker, T., & Zhang, K. (2013). Genome-wide methylation profiles reveal quantitative views of human aging rates. <i>Molecular Cell</i>, 49(2), 359-367. https://doi.org/10.1016/j.molcel.2012.10.016</p> <p>Horvath, S. (2013). DNA methylation age of human tissues and cell types. <i>Genome Biology</i>, 14(10), R115. https://doi.org/10.1186/gb-2013-14-10-r115</p> <p>Jia, T., Chu, C., Liu, Y., van Dongen, J., Papastergios, E., Armstrong, N. J., Bastin, M. E., Carrillo-Roa, T., den Braber, A., Harris, M., Jansen, R., Liu, J., Luciano, M., Ori, A. P. S., Roiz Santiañez, R., Ruggeri, B., Sarkisyan, D., Shin, J., Sungeun, K., ... Desrivieres, S. (2021). Epigenome-wide meta-analysis of blood DNA methylation and its association with subcortical volumes: Findings from the ENIGMA Epigenetics Working Group. <i>Molecular Psychiatry</i>, 26(8), Article 8. https://doi.org/10.1038/s41380-019-0605-z</p> <p>Marquand, A. F., Rezek, I., Buitelaar, J., & Beckmann, C. F. (2016). Understanding Heterogeneity in Clinical Cohorts Using Normative Models: Beyond Case-Control Studies. <i>Biological Psychiatry</i>, 80(7), 552-561. https://doi.org/10.1016/j.biopsych.2015.12.023</p> <p>Moore, L. D., Le, T., & Fan, G. (2013). DNA Methylation and Its Basic Function.</p> |

Neuropsychopharmacology, 38(1), 23-38. <https://doi.org/10.1038/npp.2012.112>

Ochana, B.-L., Nudelman, D., Cohen, D., Peretz, A., Piyanzin, S., Rosenberg, O. G., Horn, A., Loyfer, N., Varshavsky, M., Raisch, R., Shapiro, I., Friedlander, Y., Hochner, H., Glaser, B., Dor, Y., Kaplan, T., & Shemer, R. (2025). Time is encoded by methylation changes at clustered CpG sites. *Cell Reports*, 44(7). <https://doi.org/10.1016/j.celrep.2025.115958>

Reichard, J., & Zimmer-Bensch, G. (2021). The Epigenome in Neurodevelopmental Disorders. *Frontiers in Neuroscience*, 15. <https://doi.org/10.3389/fnins.2021.776809>

Ruggeri, B., Nymberg, C., Vuoksima, E., Lourdasamy, A., Wong, C. P., Carvalho, F. M., Jia, T., Cattrell, A., Macare, C., Banaschewski, T., Barker, G. J., Bokde, A. L. W., Bromberg, U., Büchel, C., Conrod, P. J., Fauth-Bühler, M., Flor, H., Frouin, V., Gallinat, J., ... Schumann, G. (2015). Association of Protein Phosphatase PPM1G With Alcohol Use Disorder and Brain Activity During Behavioral Control in a Genome-Wide Methylation Analysis. *The American Journal of Psychiatry*, 172(6), 543-552. <https://doi.org/10.1176/appi.ajp.2014.14030382>

Rutherford, S., Barkema, P., Tso, I. F., Sripada, C., Beckmann, C. F., Ruhe, H. G., & Marquand, A. F. (2023, March 13). Evidence for embracing normative modeling. *eLife*; eLife Sciences Publications Limited. <https://doi.org/10.7554/eLife.85082>

Sun, L., Zhao, T., Liang, X., Xia, M., Li, Q., Liao, X., Gong, G., Wang, Q., Pang, C., Yu, Q., Bi, Y., Chen, P., Chen, R., Chen, Y., Chen, T., Cheng, J., Cheng, Y., Cui, Z., Dai, Z., ... He, Y. (2025). Human lifespan changes in the brain's functional connectome. *Nature Neuroscience*, 28(4), 891-901. <https://doi.org/10.1038/s41593-025-01907-4>

Sun, Y., Jia, T., Barker, E. D., Chen, D., Zhang, Z., Xu, J., Chang, S., Zhou, G., Liu, Y., Tay, N., Luo, Q., Chang, X., Banaschewski, T., Bokde, A. L. W., Flor, H., Grigis, A., Garavan, H., Heinz, A., Martinot, J.-L., ... Desrivieres, S. (2022). Associations of DNA methylation with behavioral problems, grey matter volumes and negative life events across adolescence: Evidence from the longitudinal IMAGEN study. *Biological Psychiatry*, 93(4), 342-351. <https://doi.org/10.1016/j.biopsych.2022.06.012>

Are there any files you would like to upload to support your concept paper?

☒ Yes
☐ No

If yes, how many files would you like to upload?

1 ▼

File 1 - Please upload your file

[Figure 1.pdf \(0.82 MB\)](#)

Applicants: If you would like to continue your application later, please press the "Save and return later" button below. Please copy or write down the Return code provided.

To return later, you may click on "Returning?" on the top right of the screen in the E-Risk Concept Paper Form link, which is the same link that was used to access this form: <https://redcap.link/ERiskConceptPaperForm>

3. Expected project outcomes

Please note:

The stated end date must be within 24 months of the date when this form is submitted. This end date will form part of the formal data use agreement and on this date you should delete the dataset. Therefore, it must be a realistic date for

completion of the project including all analysis, writing a manuscript, review of the manuscript by all collaborators, submission, revisions, and acceptance of a paper for publication.

If you require an extension to the end date of the project, then you should contact Prof Fisher (helen.2.fisher@kcl.ac.uk) to discuss this. If you have signed a formal data use agreement, you will need to complete a form to request a licence extension. In some cases, we may also ask you to complete a new concept paper form if there have been substantial changes to the project or a long period of time has elapsed (e.g., greater than a year since the end date of the original project).

If the objective of the project is not a journal publication, please suggest an end date within 12 months instead of 24 months, and state a measurable, concrete outcome. If the objective of the project is a student dissertation, then the expected end date should be the deadline for submission of the dissertation; dissertation projects will only be accepted on agreement that they are strictly not for publication.

| | |
|--|--|
| Date form submitted | <input type="text" value="22-07-2025"/> DD-MM-YYYY |
| End date for the project | <input type="text" value="21-07-2027"/> D-M-Y DD-MM-YYYY |
| Do you expect to publish your results in a journal? | <input checked="" type="radio"/> Yes <input type="radio"/> No |
| If yes, please provide a provisional list of author names | Xinyang Yu; Sylvane Desrivieres; Di Chen; Helen Fisher; Louise Arseneault, Terrie Moffitt, Avshalom Caspi, Karen Sugden, Benjamin Williams, Eilis Hannon, Chloe Wong; Jonathan Mill; IMAGEN consortium; ENIGMA-Epigenetics Working Group |
| If yes, please provide a provisional list of journals | Nature Genetics; Nature Communications; Molecular Psychiatry; Genome Biology |

Applicants: If you would like to continue your application later, please press the "Save and return later" button below. Please copy or write down the Return code provided.

To return later, you may click on "Returning?" on the top right of the screen in the E-Risk Concept Paper Form link, which is the same link that was used to access this form: <https://redcap.link/ERiskConceptPaperForm>

4. List of variables required

Please note:

When specifying variables, please be unambiguous. For each variable, specify the name of the measure, twin age, informant, and if you want specific subscales/derived categories (e.g., Depression from interview with twin at age 18; both number of symptoms and DSM-IV diagnosis). Alternatively, for maximum clarity, give actual variable names (e.g., MDESXE18 - MDE Symptom scale - P18 - Elder; DXMDEE18 - Major depressive episode, dsm4 - P18 - Elder).

By default, the dataset will usually include twin and family IDs, the "random" and "true" twin order variables, the cohort the twin is from (1994 or 1995), twin sex, ethnicity and zygosity variables, and family socioeconomic status at age 5. These routine background variables are listed in the table below. If you require further background variables, please specify them in your list.

Access to some parts of the dataset are restricted, namely identifiable data (e.g., postcodes, video recordings, individual-level genotypic and epigenetic data) which will not be shared outside King's College London, and linked administrative data which is only accessible via the UK Longitudinal Linkage Collaboration's Trusted Research Environment (this requires a separate formal data access agreement).

Background variables that will be included by default:

| Variable name | Description |
|---------------|--------------------------|
| FAMILYID | Unique family identifier |
| ATWINID | Twin A ID (ex chkdg) |
| BTWINID | Twin B ID (ex chkdg) |
| RORDERP5 | Random Twin Order |
| TORDER | True Twin Order |
| RISKS | Sample Groups |
| COHORT | Cohort |
| SAMPSEX | Sex of Twins |
| ZYGOSITY | Zygosity |
| SETHNIC | Ethnicity of Twins |
| SESWQ35 | Social Class Composite |

Please select the variables that will be requested

- ☐ Age 5 variables
☐ Age 7 variables
☐ Age 10 variables
☐ Age 12 variables
☒ Age 18 variables
☐ Age 26 variables
☐ Age 30* variables

Age 18 variables

DNA methylation data:

Illumina 450K DNA methylation data from peripheral blood at age-18 and related variables (probes, batch number, methylation array control probe principal components, chip ID etc, cell type composition estimates) for both elder and younger twin.

Health/behaviours:

BMIE18 (BMI - P18 - Elder)

WAISTHIPE18 - Waist Hip Ratio - P18 - Elder

SMKPKYRE18 (Smoking - pack years, ages 12 to 18)

SMKCNUME18 (Smoking - current number of cigarettes)

Mental health and well-being:

DXMDEE18 (Major depressive episode, DSM-IV - age 18 - Elder)

DXGADE18 (Generalized Anxiety Disorder, DSM-IV-based - age 18 - Elder)

SHARME18 (Self-harm - age 18 - Elder)

SUICATE18 (Suicide attempt - age 18 - Elder)

DXPTSDLFE18 (PTSD lifetime diagnosis, DSM-IV - age 18 - Elder)

CDMODE18 (Moderate conduct disorder (≥ 5 symptoms) - age 18 - Elder)

PSYEXPCE18 (Psychotic experiences (categorical) - Elder)

DXADHD5X_18E (DSM-5 ADHD Dx (based on ≥5 Symp) [incl 4 NEET & meds] - P18 - ET)

DXMARJE18 (Marijuana dependency, dsm4 - P18 - Elder)

DXDRUGME18 (Drug dependent (or on methadone maintenance), dsm4 - P18 - Elder)

DXALCDEPE18 (Alcohol dependent, dsm4_based - P18 - Elder)

Are you requesting access to identifiable or linked data?

- ☐ Yes
☒ No