### A1. Project title (200 characters):

The metabolic consequences of adverse early life conditions and subsequent risk for adult cardiovascular disease and type 2 diabetes

### A2. Research question(s) and aim(s) (up to 5000 characters or 200 words):

The overarching goal for this project is to better quantify and understand the impact that early life conditions have on adult metabolic capacity and the risk for type 2 diabetes (T2D) and cardiovascular disease (CVD) in adulthood. To that aim we have two objectives:

1. To determine the association of early childhood adversity (using adult leg length as a biomarker) on various metabolic characteristics:
   * Adult liver function (using markers of fat and inflammation from MRI images).
   * Distribution of body fat, whether fat is stored primarily in the trunk or in the peripheral tissues (using DXA and bioimpediance measures).
   * Adult kidney function (using urinary markers).
   * Adult blood biomarkers (fasting glucose, lipid measures) {{confirm with what they have measured}}.
2. To determine the mediating pathways that adult metabolic capacity (using all variables listed above) has between exposure to early childhood adversity and risk for T2D and CVD.

The majority of applications to UK Biobank are for data only. As such, the first two questions we ask are whether your application involves access to samples or re-contact as this will require some additional information and as is set out in the Access Procedures (our data are not depletable, but our samples and re-contact opportunities are depletable) recontact/sample applications are assessed to a different (more exacting) standard.

Does your project require biological samples? **No**

Does your project require UK Biobank to re-contact participants: **No**

### A3. The background and scientific rationale of the proposed research project in general (up to 5000 characters or 300 words):

Early life adversity is known to increase the risk for chronic cardiometabolic diseases such as cardiovascular disease (CVD) and type 2 diabetes (T2D) in later adulthood. Early life, typically defined from conception until early childhood (~6 years), is a period characterised by substantial growth and development. Extreme periods of or consistent exposure to adversity, such as poverty, malnutrition, or famine, can set individuals on a higher disease risk trajectory. For instance, exposure to famine during early life can increase the risk of developing T2D by about 60%. While extreme early life adversity is relatively uncommon in contemporary European populations, some immigrants and refugees have been exposed to more extreme conditions (e.g. refugees from conflict-ridden Syria). Even less extreme adversity, such as neglect, loss of a parent, or childhood conflict, can increase the risk for cardiometabolic disease. For instance, neglect and emotional abuse can increase the risk of obesity, a strong predictor of CVD and T2D, by up to 50%. These forms of adversity are much more common in many EU countries, with nearly 30% of EU children experiencing some form of mental or physical abuse.

A crucial step to developing evidenced-based solutions to mitigate risk is in understanding the pathophysiological mechanisms underlying the association between early life adversity and the development of cardiometabolic disease. Limited data exists that have investigated this area. A few studies suggest that lower early life socioeconomic position and other indirect indicators of early life conditions associates with a lower metabolic capacity. There is very limited knowledge on how early life adversity mediates its effect on later disease through adult metabolic capacity (such as through lipid metabolism, inflammatory processes, and/or glucose regulation). This project’s aim is to investigate potential metabolic changes that result from early life adversity and the resulting risk for cardiometabolic disease.

### A4. A brief description of the method(s) to be used (up to 5000 characters or 300 words):

Words: 223 / 300

For objective 1, we will analyse the data in two stages. First, we will run individual regression models for each of the variables, adjusted for potential confounders, with each metabolic variable as the outcome and early life conditions as the predictors. The next step will be to run supervised dimensionality reduction and clustering techniques, with early life conditions as the supervising set. These techniques will, together, provide a broad view of how early life conditions associate with adult metabolic function.

For objective 2, we will use a causal structure learning algorithm to identify the contributions of individual metabolic variables (using multiple aspects of metabolic function from the various metabolic measurements) that may mediate the association between early life adversity and the development of cardiometabolic disease. The algorithm constructs causal pathways in an iterative approach, while also taking into account the inherent interdependence of the metabolic variables. The strengths of this causal structure learning method over others are a) the ability to make use of high dimensional metabolic data, b) to estimate the most probable directions of the pathways, and c) to identify the most likely causal structure between exposure (early life adversity), metabolic variables, and outcome (CVD or T2D). This approach is necessary as the method must consider interdependence between variables, to quantify the magnitude of association, and to determine the direction of the pathway.

### A5. The type and size of dataset required (e.g., case-control subset, men only, imaging data only, whole cohort, etc.) (up to 5000 characters or 100 words):

Whole cohort, MRI imaging, and the biomarker data.

### A6. The expected value of the research (taking into account the public interest requirement) (up to 5000 characters or 100 words):

Word count: 103 / 100

We expect our results to provide etiological insight into how early life exposure to adversity modifies adult metabolic capacity, which may subsequently reveal specific metabolic patterns (e.g. more circulating lipids or greater storage of fat in liver) or a greater fat disposition in abdominal adipose tissue. These changes in metabolic capacity could be used as targets for individual level prevention or management in those exposed to adversity. We also intend to publish the analytic methods in a format easily usable by other researchers, e.g. as an R (statistical programming language) software package, which we hope will accelerate research in this area further.

### A7. Please provide up to 6 keywords which best summarise your proposed research project:

early life conditions, etiology of type 2 diabetes, etiology of cardiovascular disease, metabolic capacity, causal structure learning analysis, multivariate analysis

### A8. Please provide a lay summary of your research project in plain English, stating the aims, scientific rationale, project duration and public health impact (up to 5000 characters or 400 words):

Word count: 195 / 400

We aim to study how early life adversity may influence adult metabolism and eventually lead to diseases such as cardiovascular disease and type 2 diabetes. While we know that early life adversity negatively affects health in later adulthood, we don’t know how it may affect metabolism. Historically, this area of research has been difficult to study due to how hard it is to get valuable and suitable data, but also because analysing the data is incredibly complicated. With recent advances in machine learning and big data, better tools are becoming available that can begin answering how early life conditions influence later adult metabolic capacity. Over the course of 48 months, we not only learn more about the “how” of early life adversity’s role in disease, but also aim to create better tools for the analyses so other researchers can further accelerate their own work. In the end, we hope that this knowledge will help with creating more effective public health strategies for disease prevention and with making more precise clinical decisions for disease management in those exposed to early life adversity.

### A9. Will the research project result in the generation of any new data fields derived from existing complex datasets, such as imaging, accelerometry, electrocardiographic, linked healthcare data, etc, which might be of significant utility to other researchers:

*No*

### A10. What is the estimated duration of your project, in months? If you consider (because for example the project is one involving the generation of hypotheses) that it would be difficult to set a fixed end point, we are prepared to consider a rolling 3-year period (during which annual updates are required):

*48 months*

Please note that you are expected to publish (or to make publicly available) your results and return to UK Biobank:

* any important derived variables
* a description of the methods used to generate them
* the underlying syntax/code used to generate the main results of the paper, and
* a short layman’s description that summarises your findings.

These should be provided within six months of each publication or within 12 months of the project end date (whichever comes first). We also ask that you send us a copy of your accepted manuscript at least two weeks prior to publication and alert us if there are any ethical or contentious issues surrounding the findings.

### B. Selection of data-fields

* Blood counts
* Urine assays
* Body size measures
* Abdominal MRI (liver images)
* DXA (region specific bone and body composition)
* Impediance measures
* Early life factors
* Indices of Multiple Deprivation
* Genotype/genes