**Twin Study Application Questionnaire:**

1. **Scientific Hypothesis** (please include a **detailed and comprehensive** description)

Control Hypothesis with Twins Data:

"To what extent does genetic similarity among monozygotic twins, as compared to dizygotic twins, account for the variability observed in blood biomarker levels, independent of environmental factors?"

MSc Hypothesis:

"Blood biomarkers display measurable heritability, with certain biomarkers showing stronger genetic associations."

The central hypothesis of this research is that genetic variability significantly influences blood biomarker levels, independent of environmental factors, lifestyle choices, or other external influences. By examining twin pairs, as my control—particularly monozygotic (identical) twins who share nearly identical genetic profiles—the study aims to isolate genetic contributions to blood biomarker variability. This approach enables a comparison between monozygotic and dizygotic (non-identical) twins to gauge the extent to which blood biomarker levels are heritable versus influenced by environmental or lifestyle factors.

Once I complete the twin study as my control, I will apply these findings to the UK Biobank dataset to verify if the twin results align with trends in the general UK population. The twin studies are crucial for establishing my initial findings. However, if I discover that genetic influence has minimal impact on blood biomarker outcomes, I will need to reconsider my hypothesis.

1. **Scientific Rationale** – Define the aims of the proposed research including the research question(s) you are aiming to answer

The aim of this research is to assess the extent to which genetic variability influences blood biomarker levels. By studying twin pairs, particularly monozygotic twins, this research seeks to isolate genetic factors from environmental and lifestyle influences to determine the heritability of key biomarkers. The primary research questions are:

How much of the variability in blood biomarkers can be attributed to genetic factors?

Are there specific genetic markers (SNPs or polygenic scores) that consistently affect certain biomarkers?

Can these genetic findings from the twin study be applied to the general population in the UK Biobank?

This study will establish a foundation for developing a genetically-adjusted baseline model to improve disease prediction accuracy.

1. **Data/Material Requirement** - Please include a description of the data and/or the quantity and type of samples required

This research requires a comprehensive dataset from the twin studies, including complete genetic variability information through SNP data and polygenic risk scores. These genetic data points are essential to identify specific genes or larger gene networks impacting blood biomarker levels. To ensure compatibility with the UK Biobank dataset, I will also need lifestyle and environmental phenotype data from the twin studies, closely matching those available in the Biobank.

Although the Biobank contains more extensive phenotyping, the twin studies include some unique features not available in the Biobank. Therefore, having a full dataset covering both genomic and phenotypic data across both resources is critical for aligning findings. This alignment will help ensure that results from the twin study can be reliably compared and applied to the broader Biobank population. Due to the initial uncertainty around the significance of specific markers, access to a wide range of genetic and phenotypic markers will allow for credible, accurate conclusions.

1. **Methodology/Analysis Plan** for the data/samples required – Please provide a non-technical description of how the research will be undertaken.

The research will begin with twin data to establish key categories that may influence blood biomarker variability. The data will be organized into categories such as:

1. Genetic:

- SNPs: Single nucleotide polymorphisms linked to biomarkers.

- Polygenic Scores: Combined influence of multiple SNPs on traits.

- Other Genetic Categories: Potential novel or less-studied

genetic markers that may influence biomarkers.

- Other

2. Environmental:

1- Location: Regional and environmental exposures.

3. Measured Factors:

- Demographics: Age, gender.

- Physical Health Metrics: Weight, height, blood pressure, and body mass index (BMI).

4. Blood Phenotype Variables:

- Timing and Conditions: Time and volume of blood draw.

- Biomarker Ranges: Method for aligning and categorizing biomarkers with minor differences (e.g., 5.3 vs. 5.7).

- Key Biomarkers: Identification of biomarkers frequently used in clinical disease prediction.

5. Mental Health (if available):

Psychological Factors: Stress levels or other mental health metrics.

6. Dietary and Lifestyle Factors:

- Diet: Types of food, nutritional intake.

- Lifestyle: Activity levels, habits.

7. To maintain consistency, I will filter out participants with known factors that may skew blood biomarker data, such as:

- Chronic diseases that influence biomarkers.

- Specific diets that may impact blood markers.

- Medications known to alter biomarker levels.

With these filtered categories, I will conduct analyses comparing each set of variables to blood biomarker levels:

1. Blood markers vs. Genetics (SNPs and polygenic scores).

2. Blood markers vs. Environment.

3. Blood markers vs. Measured Demographics.

4. Blood markers vs. Mental Health.

5. Blood markers vs. Diet/Lifestyle.

For each category, I will create an influence score reflecting its impact on biomarker variability, ultimately combining scores to estimate a percentage of influence on biomarkers.

To continue with the study, I’ll establish a threshold where genetic influence must account for at least 50% of blood biomarker variability to proceed with modeling genetic adjustments for broader population predictions. This threshold is scientifically reasonable, as it suggests a strong genetic component that is significant enough to justify further investigation into genetic contributions to biomarker baselines.

1. **Supporting Research** Please provide details of any pilot studies undertaken or experiments and design that will support this application

Currently, no pilot studies or preliminary experiments have been undertaken for this project. Under the guidance of my supervisor, Dr. Manuel Corpus, I am laying the groundwork for a research plan that may extend into a PhD. The aim of this extended research is to adjust blood biomarker levels based on genetic profiles to enable more precise diagnostic and treatment tools in various clinical settings.

In the PhD phase, I intend to apply findings on genetic influence over blood biomarkers to improve accuracy in clinical biomarker-based assessments, potentially enhancing diagnosis and treatment strategies. I also plan to test adjusted biomarker models on AI platforms, such as MITON, to evaluate if genetic adjustments improve disease predictability scores.

As a lung and liver transplant survivor with diabetes, cystic fibrosis, and mental health conditions, I’ve personally experienced the limitations of current biomarker tests, which don’t always account for individual genetic differences. Given that billions of people rely on standardized biomarker tests, adjusting these tests for genetic differences holds significant potential for improving patient care, moving beyond the limitations of one-size-fits-all biomarker assessments.