Project Title: Genetic Influence on Blood Biomarkers: A Twin Study Approach

Application Id	355
Status	New
Date Submitted	
Date Approved	

Main Applicant: Mr Donald Philp

Email	w2107579@westminster.ac.uk
Tel	07778904717
Position	Postgraduate Student
Department	Life Science
Institution	University of Westminster
Country	England
ORCID iD	
Address	1a Dawlish Road London NW2 4HP England

Co Applicant: Dr Manuel Corpus

Email	M.Corpas@westminster.ac.uk
Tel	
Position	Course Leader
Department	Life Sciences
Institution	University of Westminster
Country	England

Project Details

Title	Genetic Influence on Blood Biomarkers: A Twin Study Approach
Proposed Start Date	01/10/2024
Proposed End Date	01/12/2025

Funding and Costs

Please note that applications for funding must be reviewed by the DTR Committee prior to submission to a funding body and may require a DTR member to be a co-applicant. Applications must be submitted for review by our committee at least three weeks before the funder submission deadline.

All researchers accessing TwinsUK data/biological samples will be charged on a cost recovery basis. This cost will vary depending on the amount and type of data/biological samples.

Source of Funding	Self funding

Ethical Approval

Does the study supporting your project have ethical approval from an Ethics Committeee or an Institutional Review Board?	Not Applicable
You intend to carry out analysis of existing data where generic DTR ethics approval will operate. Please provide details.	This study will adhere to strict ethical standards to protect participant privacy and uphold data security. Key ethical considerations include: 1. Informed Consent: All participants in the twin study and Biobank have provided informed consent for the use of their data in scientific research. The research will comply with the consent agreements that outline permitted uses, ensuring participants are respected and fully informed about how their data contributes to the study. 2. Data Privacy and Confidentiality: We will implement secure data handling practices in line with GDPR (General Data Protection Regulation) and other relevant standards. Personal identifiers will be removed to anonymize the data, ensuring that individual privacy is protected. 3. Data Security: Secure servers and encrypted data storage methods will be used to prevent unauthorized access to sensitive information. Only authorized personnel will have access to the dataset, and data sharing will follow Biobank and twin study policies. 4. Minimizing Harm: The study is observational, with no direct interventions. Thus, it poses minimal risk to participants. Analyses are based on pre-existing data, so there is no physical interaction with participants, reducing any potential harm. 5. Transparency and Accountability: Findings will be reported transparently and in a way that contributes to scientific understanding while ensuring no misuse of genetic information. Data usage will be monitored, and ethical review boards will be consulted to ensure ongoing compliance with ethical standards throughout the research process. 6. Benefit to Society: The goal of this study is to advance knowledge on the genetic influence on blood biomarkers, potentially contributing to improved personalized medicine and public health. This aligns with broader ethical objectives to use scientific research for societal benefit.

Data Requirement

We have extensive clinical, physiological, and behavioural and lifestyle data, including biochemical and genetic data, available to researchers, as well as hundreds of phenotypes related to common diseases. Many data types have been collected at multiple longitudinal time points. Available sample types include: DNA, cell lines, serum, plasma, stool and urine. Please tick below category of data and/or sample you require:

Is existing data required?	Yes
Are existing biological samples required?	No
Is new data required?	No
Are new biological samples required?	No

Existing Data

Please indicate the data/variables you would like to request to answer your proposal hypothesis:

- Data from twin visits & Questionnaires
- Genotype data Raw
- Genotype data Imputed
- Epigenetic data (Methylation)
- Expression data (Transcriptomics)
- Metabolomic dat

Twins set required (if subset please provide full description on the Scientific outline)	All Twins

The available phenotypes data can be found in the Phenotype Spreadsheet December 2023 Please download this spreadsheet, complete the “Data Request” tab and upload the completed request in the Supporting Documents section at the end of this form.

Phenotypes Domain	1. Genetic Markers: Focus on SNPs, DNA analysis, and polygenic scores that impact biomarkers. 2. Blood Biomarkers: Includes cholesterol, glucose, insulin, and related blood-based markers. 3. Dietary & Lifestyle Factors: Involves dietary habits, alcohol consumption, smoking, and general lifestyle factors. 4. Mental Health & Psychological Factors: Includes markers related to stress, anxiety, and other mental health factors. 5. Environmental Exposures: Environmental and regional factors that could impact blood biomarker variability. 6. Physical Health Metrics: Metrics like weight, height, BMI, and blood pressure. 7. Reproductive Health: Hormonal factors, contraceptive use, pregnancy history, and menstrual health. 8. Family History: History of chronic diseases, specific traits, or hereditary conditions among family members. 9. Chronic Disease History: History of conditions such as diabetes, hypertension, and cardiovascular diseases. 10. Medication & Treatment: Medication use, including prescriptions, chemotherapy, or hormone replacement therapy.
Number Phenotypes	5519
Will you require help from DTR statisticians for data analysis?	No

Scientific Proposal

Please ensure that your proposal **clearly highlights** the specific project data requirements stating the hypothesis, rationale for using this data, analysis plan, data specification, and other study methods considered.

Please provide **extensive and comprehensive** information when filling in the fields below. **Lack of information will impair the committee ability to decide your application outcome.** Ensure that you list the phenotype domain(s) or variables summary you request, and that the final list of phenotypes is attached at the end of this form. The phenotype data must include the DTR variable codes of interest, either P (phenotype) or Q (questionnaire). Please see the list of phenotype and variables in the Phenotype Data Spreadsheet.

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.
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.

Scientific Proposal

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.

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.

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.

Agreement

Agreement

By signing this form, I confirm that I understand and agree to comply with the conditions stipulated below.

The Department of Twin Research in accordance with King's College London policy will not be permitted to release new and /or identifiable data/samples until a Material Transfer Agreement (MTA) has been finalized for identifiable material or a Data Transfer Agreement (DTA) has been finalized for new (non-identifiable) data.

1. The data may only be used for non-commercial academic research. The data and the results of the research may not be used for commercial purposes unless a revenue-sharing agreement or commercial license is drafted and processed by King's College London Business.
2. No data will be passed to third parties or journals without written permission from the Department of Twin Research.
3. The data remains the property of King's College London and if any new variables are derived from the data and /or any changes are made to the data, these will be returned to the Department of Twin Research upon acceptance for publication by a Journal or at the latest within six months from the end of the project, and any new variables derived from the data and/or changes made to the data shall be the property of King's College London.
4. No attempt should be made to link or combine the data provided under this agreement to other information or archived data available for the data sets provided, even if access to that data has been formally granted to you, or it is freely available without restriction, unless specific permission to do so has been received from the relevant access committee(s) or sample custodians.
5. The Department of Twin Research and its funder's contribution to this project will be acknowledged in any resulting publications or dissemination material.
6. All manuscripts and drafts of oral presentations will be submitted to the Department of Twin Research for review and approval at least 15 days before submission or presentation. A final version of the manuscript and summary of any oral presentations will be sent to the department on final submission.
7. Authorship will be agreed by mutual consent. All publications will have to acknowledge the TwinsUK resource. Standard acknowledgements are available at https://twinsuk.ac.uk/
8. The identity of the twins should be protected at all times and no contact or tracing attempts will be made.
9. The Department of Twin Research conforms to GDPR standards. By signing this document, you confirm that your organisation conforms to GDPR standards and the data you receive from us will be regulated by the GDPR.

I confirm that I understand and agree to comply with the conditions stipulated above	No (Cannot submit application without this)
I have discussed this application with all co-applicants, and/or my supervisor and they understand and agree to comply with the conditions stipulated above (Cannot submit application without this)	**No (Cannot submit application without this)**

Supporting Documents

File Name	File Description
Phenotypes_Application.xlsx	Phenotype selected

Final Checks

Confirm phenotype request sheet attached?	No

Document Produced: 04/11/2024 11:38:40