Centre for Longitudinal Studies

CLS register of approved GENDAC projects

Reference	GDAC_2026_12
Project title	Testing a UCL genomics pipeline to generate polygenic scores
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Many human characteristics, from physical traits to aspects of health and behaviour, are shaped partly by genetics, with thousands of tiny genetic differences each adding a small effect. A polygenic index combines these differences into a single score estimating a person's genetic likelihood of having a particular characteristic. We are building a pipeline to calculate these scores for participants in major UK cohort studies Including the Millennium cohort study and other cohorts held by the centre for longitudinal studies. To know whether the scores work, we need to check them against reality: do people with a higher score for a trait actually tend to show that trait? We will compare each calculated score with the information the studies have already collected from participants. We use only genetic findings already published by other researchers. This work helps the research community judge how reliably such scores predict real-world outcomes.
Keywords	Cohorts, genomics, MCS, pipeline, and polygenic scores
Date received	12 June 2026
Date approved	17 July 2026

Reference	GDAC_2026_11
Project title	Genetic pathways linking parental mental health and adolescent suicidality
Institution	McGill University (Canada)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Suicide is a leading cause of death among adolescents worldwide. A parental history of psychopathology is associated with an increased risk of suicidality in offspring. However, it remains unclear whether mental health problems in mothers and fathers carry similar risks for their children's self-harm or suicide attempts, and whether genetic factors help explain these links. This project aims to quantify and explain the links between parental mental health and adolescent suicidality (self-harm, suicide attempt). First, we will estimate and compare the magnitude of associations between maternal versus paternal mental health and adolescent self-harm and suicide attempt. Second, we will

	quantify the proportion of these associations explained by genetic liability. These analyses will help clarify sex-specific pathways linking parental mental health to adolescent suicidality and identify family-level targets for prevention.
Keywords	Family history, gender difference, mental health, polygenic risk, and suicide attempt
Date received	24 March 2026
Date approved	6 July 2026

Reference	**GDAC_2026_10**
Project title	PEpi Score: Polygenic Epistatic Scoring
Institution	University of Edinburgh
Application type	Data access / Genetics
Cohort	NCDS, BCS70, Next Steps, and MCS
Abstract	Human genetic studies generally examine the association between the outcome and (a) single genetic markers and (b) an additive collection of multiple genetic markers. However, epistasis, or gene-gene interactions, are commonly found in mammalian animal models, and they can explain a sizeable proportion of the variation in complex traits. This study seeks to develop a novel statistical method that models the additive portion plus the epistatic component of several human traits.
Keywords	Epistasis, Gene-gene interaction, GWAS, meta-analysis, polygenic scores, and simulation
Date received	2 February 2026
Date approved	12 May 2026

Reference	**GDAC_2026_09**
Project title	Harnessing the Uks birth cohorts to detect preclinical dementia in the population (Insight58/70)
Institution	University College London (UCL)
Application type	Samples request / Blood
Cohort	NCDS and BCS70
Abstract	This project follows on from an ARUK funded study – Insight46, which was a neuroscience sub study of the 1946 birth cohort, comprising of cognitive testing, MRI/PET imaging, cardiac and biomarker assessments. This project will extend this work into the next two birth cohorts – the 1958 National Child Development Study (NCDS) and the 1970 British Cohort Study (BCS70). The project aims to assess the risk of preclinical AD/dementia from blood screening, genetics and cognitive testing. Blood screening will be conducted using assays which have already been collected as part of previous surveys (assays will be transferred from where they are stored in Bristol to a UCL laboratory for analysis). The cognitive tests will be administered using 'Cognitron' and will be sent to study members to complete following their participation in the 2026 NCDS-BCS70 Survey (those who do not take part in the survey will

	still be invited to complete cognitron). A sub-sample of c. 100 participants will then be invited to a clinic visit for MRI/PET imaging.
Keywords	Alzheimer's, biomarkers, birth cohort, cognition, dementia, and life course
Date received	24 February 2026
Date approved	24 February 2026

Reference	**GDAC_2026_08**
Project title	Multi-cohort meta-analysis of impulsivity and risk-taking
Institution	University of Edinburgh
Application type	Data access / Genetics
Cohort	NSHD, NCDS, BCS70, Next Steps, and MCS
Abstract	We will perform a multi-cohort meta-analysis of impulsivity and risk-taking, including bivariate mixture models for polygenic signals between the two using MIXeR, sample sizes permitting. We will consider these two outcomes (found in different cohort studies) and assess whether the genomic data suggest they share a similar genetic component (via polygenic scores and genetic correlations, as well as significantly associated genes/regions) and consider whether they should be combined as a meta-analysis. Of course, the studies that examine risk and others that consider impulsivity may likely be combined via meta-analysis. Sex and PCAs will be included in the models as potential confounders. Both SNPs and polygenic scores will be considered as exposures.
Keywords	Gene, GWAS, impulsivity, meta-analysis, and risk-taking
Date received	23 February 2026
Date approved	12 May 2026

Reference	**GDAC_2026_07**
Project title	Assessment of genome-wide and polygenic association with schizotypy across multiple cohorts.
Institution	University of Edinburgh
Application type	Data access / Genetics
Cohort	NSHD, NCDS, and BCS70
Abstract	Schizotypy or sub-clinical experiences of psychosis are constructs that encompasses a set of personality traits bearing a resemblance to the clinical manifestation of schizophrenia. This association is thought to be underpinned by common underlying genetics, as higher levels of schizotypy-like experiences are most frequently found in individuals with a family history of schizophrenia. However, unlike schizophrenia, studies of schizotypy-like experiences are not systematically confounded by the influences of active psychosis or anti-psychotic medication. Therefore, it may be an important intermediate phenotype as provides a framework for investigating the etiology and pathophysiology of schizophrenia spectrum disorders and thus improve understanding surrounding the genetic, environmental and psychological architecture. We sought to understand the genomics of schizotypy-like experiences

	using multiple cohorts and performing a harmonization step across studies to combine questionnaire features to perform a meta-analysis to increase statistical power.
Keywords	Gene, GWAS, meta-analyses, polygenic, schizophrenia, and schizotypy
Date received	23 February 2026
Date approved	12 May 2026

Reference	**GDAC_2026_06**
Project title	Exploring the causes of adolescent mental health problems with consideration of genetics
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	MCS
Abstract	Mental health problems affect many children and adolescents and can have long-lasting impacts on their well-being. This study aims to understand how both adverse childhood experience (ACEs), such as child maltreatment, and lifestyle behaviours, such as physical activity and sleep influence mental health and to examine the role that genetics plays in these associations. To do this we will use self-reported and genetic data from the UK Millennium Cohort Study. Exposures include ACEs and lifestyle behaviours. Outcomes include emotional and behavioural difficulties, psychological distress, and self-harm. We will conduct genetic sensitivity analyses by using genetic scores to evaluate whether someone's genetic liability to poor mental health explains associations between ACEs, lifestyle behaviours and mental health. The findings will help identify key risk factors and inform strategies for prevention and early intervention, especially for young people who may be genetically vulnerable to mental health problems.
Keywords	Childhood adversities, genetic confounding, genetic sensitivity analysis, mental health, polygenic scores, and risk factors
Date received	16 February 2026
Date approved	13 April 2026

Reference	**GDAC_2026_05**
Project title	Multi-cohort meta-analysis of the genomics of perceived and observed social support
Institution	University of Edinburgh
Application type	Data access / Genetics
Cohort	NSHD, NCDS, BCS70, Next Steps, and MCS
Abstract	Genes play an important role in shaping social behaviour as they modulate brain activity through molecular pathways; therefore, it can be said that genes regulate the expression of behaviour. Social support is correlated with an individual's mental health as these social interactions require effective communication and participation. It may be that how someone perceives their social support or how frequently they participate in social support have different levels of genetic control

	and/or have different genetic pathways. We would like to use information from these cohorts to perform a genetic analysis of social support by studying differences in the DNA of individuals, harmonise these data with data from different cohorts and perform an overall study combining the individual study results.
Keywords	Genes, GWAS, meta-analysis, observed social support, and perceived social support
Date received	20 January 2026
Date approved	30 April 2026

Reference	**GDAC_2026_04**
Project title	Interactions between genetics and obstetric complications of mental ill health
Institution	University of Edinburgh
Application type	Data access / Genetics
Cohort	NSHD, NCDS, BCS70, and Next Steps
Abstract	Genetic interactions with obstetric complications (OCs) such as those producing a lack of oxygen to the baby, have long been hypothesised to influence later mental ill health. However, this has not often been directly tested. Accumulating evidence suggests OCs impact future health and they have been implicated in the aetiology of psychiatric disorders. OCs are hypothesised to interact with underlying genetics to increase risk. I will undertake a genome-wide search of interactions with OCs for a variety of mental health outcomes in multiple cohorts and perform a meta-analysis of the cohorts together to increase power. Defining these interactions could lead to early intervention for individuals experiencing obstetric complications who carry the SNP at hand. It will also contribute to better understanding of the aetiology of these measures of mental ill health.
Keywords	Extraversion, gene-by-environment interaction, GHQ, GWAS, meta-analysis, MDQ, neuroticism, obstetric complication, and schizotypy
Date received	20 January 2026
Date approved	30 April 2026

Reference	**GDAC_2026_03**
Project title	The contribution of language/communication, cognition and social abilities to life-course development and (mental) health trajectories: a genetic investigation.
Institution	Max Planck Institute for Psycholinguistics (The Netherlands)
Application type	Data access / Genetics
Cohort	BCS70
Abstract	Language, communication, cognition, and social behaviour are core developmental abilities that enable children to participate in society, shaping later social, educational and health outcomes. Individual differences in these domains emerge early in life, are partly heritable, and partly influenced by environmental exposures, yet the mechanisms

	underlying developmental variation, risk, and resilience, in the presence of early-life stress remain poorly understood. This project aims to investigate the genetic and environmental contributions to language social behaviour, and related cognitive and emotional outcomes across childhood and adolescence, and to examine how early developmental trajectories relate to later-life outcomes. Using longitudinal, multi-informant, and genomic data, the project will: (i) characterise the genetic architecture of early language, communication, and social abilities; (ii) examine links between childhood language and social behaviour and later mental health, educational, and social outcomes; and (iii) investigate genetic and environmental contributions to resilience following early-life stress.
Keywords	Cognitive skills, early-life exposures, education, genetics, health, home environment, language, life-course development, mental health, social behaviour, and social communication
Date received	16 January 2026
Date approved	25 March 2026

Reference	**GDAC_2026_02**
Project title	The role of genetic susceptibility in moderating the impact of environmental risk factors on the developmental trajectories of depression during childhood and adolescence.
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	The availability of large-scale genetic data has renewed interest in how genes and the environment interact to shape mental health. This project will examine how genetic risk for depression influences the impact of life challenges during childhood and adolescence. Research shows that being overweight, facing peer difficulties, and experiencing bullying are linked to higher depression levels at age 14 in both genders. In girls, additional factors like low family income, parental mental health issues, and higher childhood cognitive scores also affect well-being. This study will explore whether individuals with a higher genetic risk for depression are more susceptible to developing the condition when exposed to multiple stressors and whether these effects vary by gender. Using diverse research methods and longitudinal data provided by Millennium Cohort Study, this project aims to inform early intervention strategies for adolescent depression.
Keywords	Adolescent mental health, gene-environment interaction, genetic risk and depression, longitudinal mental health Research, and mental health vulnerability
Date received	9 January 2026
Date approved	12 May 2026

Reference	**GDAC_2026_01**

Project title	Exploring the relationship between antiseizure medication use in pregnancy and offspring birthweight
Institution	University of Melbourne (Australia)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Taking antiseizure medications (ASMs) during pregnancy has recently been associated with an increased risk of having a baby whose birthweight is lower than normal for the number of weeks of pregnancy. However, whether this association is truly due to ASMs or to the condition for which ASMs are prescribed, such as epilepsy, is unclear. Single nucleotide polymorphisms are small changes in DNA, which individually have a small effect, but collectively can predispose someone to certain traits, including having low birthweight. This genetic predisposition to a trait is measured with a polygenic risk score (PRS) Our study aims to compare the PRS for birthweight in ASM-exposed children with unexposed children. We hypothesise that the lower-than-normal birthweight in babies exposed to ASMs during pregnancy is not explained by the usual genetic factors influencing birthweight in the general population; rather, other factors are likely implicated, such as ASMs themselves or even a genetic predisposition to ASM-associated low birthweight.
Keywords	Antiseizure medications, epilepsy, pregnancy, small-for-gestational age, and teratogenicity
Date received	2 January 2026
Date approved	26 January 2026

Reference	**GDAC_2025_14**
Project title	Dissecting Risk Factors and Comorbidities in Chronic Respiratory Diseases: An Integrative Epidemiological, Environmental, and Genetic Approach
Institution	Guangzhou Medical University (China)
Application type	Data access / Genetics
Cohort	NCDS, BCS70, Next Steps, and MCS
Abstract	Long-term breathing conditions like asthma and COPD affect millions of people. A key puzzle is why some individuals develop these conditions, and why they often suffer from more than one at once, making life more difficult. Our study asks a fundamental question: How do our genetics, lifestyle, and environment work together to influence our risk of developing these lung diseases? By carefully studying the valuable health information that you and other cohort members have provided over your lifetimes, we can search for important clues and risk factors. Our goal is to piece together this puzzle. Understanding these connections could lead to better ways to prevent and treat these common but serious conditions, ultimately helping people lead healthier lives.

Keywords	Asthma, chronic respiratory disease, comorbidity, COPD, epidemiology, genetic association, risk factors, and UK Biobank
Date received	24 November 2025
Date approved	13 April 2026

Reference	**GDAC_2025_13**
Project title	Deriving DNA methylation scores in the NCDS and BCS70
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	NCDS and BCS70
Abstract	This project will create new data on DNA methylation – chemical markers on our DNA that help control how genes work – for participants in two long-running UK studies: the 1958 National Child Development Study (NCDS) and the 1970 British Cohort Study (BCS70). DNA methylation does not change the DNA sequence itself, but it can reflect early-life experiences, health, ageing, and environmental influences such as smoking or stress. By generating these measures, we will give researchers a valuable new resource for studying how life experiences "get under the skin" to influence health and wellbeing across adulthood. The new data will be linked securely to existing survey and health information already collected in these studies, with strict safeguards to protect participant privacy. This resource will support future research on topics such as ageing, mental and physical health, and the long-term effects of childhood environments.
Keywords	Cohorts, DNA, health, epigenetics, and methylation
Date received	3 December 2025
Date approved	3 December 2025

Reference	**GDAC_2025_12**
Project title	Quantifying the Contribution of Air Pollution to Socioeconomic Gaps in Child and Adolescent Mental Health: A Millennium Cohort Study Analysis
Institution	Imperial College London
Application type	Data access / Genetics
Cohort	MCS
Abstract	This study explores how exposure to air pollution and other environmental factors (urban noise and (lack of) greenspace) in early life might contribute to mental ill health, neurodevelopmental disorders, and related behaviours in adolescence. Unlike previous work that focuses on cognitive development, we plan to use data from the Millennium Cohort Study to apply advanced causal methods to examine whether these environmental pollutants explain part of the association between socioeconomic disadvantage and mental health outcomes. We will also test how sensitive the results are to different assumptions, for example, using alternative buffer sizes around homes, accounting for noise and green space, and exploring early-life versus cumulative pollution exposure.

	In the second part of the project, we will explore whether a child's genetic makeup, measured using polygenic risk scores, may make them more sensitive to the effects of pollution. This research aims to offer a novel approach to better understanding complex pathways how environmental, biological, and social factors come together to influence young people's mental health and development.
Keywords	ADHD, ASD, air pollution, causal mediation, environmental risk, gene–environment interaction, mental health, methodology, neurodevelopment, polygenic risk score, and socioeconomic status
Date received	17 November 2025
Date approved	26 January 2026

Reference	**GDAC_2025_11**
Project title	Age of Infection: identifying novel infectious causes of non-communicable disease
Institution	University of Bristol
Application type	Samples request / Blood
Cohort	NCDS and BCS70
Abstract	Many common diseases in adulthood, like cancer or heart disease, may be caused by infections people had when they were young. We think when a person gets an infection (their 'age of infection') is a critical factor for their future health risk. However, it's currently very difficult to find out when people were first infected with different bugs. This project will use small amounts of stored blood (serum or plasma) from cohort members, collected at different ages. We will test these samples for antibodies to many common infections. This will help us create a timeline, or 'atlas', showing when people typically get these infections.
Keywords	Age of acquisition, birth cohort, epidemiology, infection, non-communicable disease, and serology
Date received	20 October 2025
Date approved	17 March 2026

Reference	**GDAC_2025_10**
Project title	Internalising and externalising symptoms in children and adolescents conceived using assisted reproductive technology: an individual participant meta-analysis
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	This research seeks to address the following questions: How do child obesity and psychological disorders develop, and what is the nature of their interrelationship? What impact do parental investments, both material and in terms of time, have on child obesity and mental health disorders? Additionally, the study will explore how obesity and mental health disorders influence children's health, cognitive abilities, social-

	emotional skills, and academic performance across various developmental stages. Obesity and psychological disorders are the two most prevalent health issues among children. Knowing their interrelationship is crucial for devising strategies that improve children's well-being. The combination of the MCS survey data and genetic data will offer valuable information on the background of the participants and the dynamic relationship between obesity and mental disorders. The researchers will estimate what a statistical model of how different inputs translate into outputs, known in the economic literature as production function. This research has a potential to offer policy recommendations for interventions aimed at enhancing the health and educational achievements of children.
Keywords	Assisted reproductive technology, externalising symptoms, infertility, internalising symptoms, and meta-analysis
Date received	7 November 2025
Date approved	26 January 2026

Reference	**GDAC_2025_09**
Project title	From Health to Homework: Exploring the Dynamics of Obesity, Mental Health, and Academic Achievement in Youth
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	This research seeks to address the following questions: How do child obesity and psychological disorders develop, and what is the nature of their interrelationship? What impact do parental investments, both material and in terms of time, have on child obesity and mental health disorders? Additionally, the study will explore how obesity and mental health disorders influence children's health, cognitive abilities, social-emotional skills, and academic performance across various developmental stages. Obesity and psychological disorders are the two most prevalent health issues among children. Knowing their origins is crucial for devising strategies that improve children's well-being. The combination of CLS survey data and genetic data will offer valuable information on the background of the participants and the development of obesity/mental disorders. The researchers will use economic methods to analyse the data and build a model to explain the findings. This research has a potential to offer policy recommendations for interventions aimed at enhancing the health and educational achievements of children.
Keywords	Human capital, mental health, obesity, and parental investment
Date received	20 October 2025
Date approved	25 March 2026

Reference	**GDAC_2025_08**

Project title	Genes, Skills, and the Intergenerational Transmission of Human Capital
Institution	University of Cambridge
Application type	Data access / Genetics
Cohort	MCS
Abstract	This project studies how family background and genetic differences jointly shape education, jobs, and income across generations. Using large UK studies that combine survey data and genetic information, we focus on two types of genetic predispositions: one related to educational attainment and another related to mental health. We ask how these predispositions, together with parents' time and financial support, influence children's skill development, career choices, and long-term economic outcomes. Our work also examines how advantages and disadvantages are transmitted across generations, and whether supportive environments can reduce inequalities linked to genetic differences. By combining evidence from both childhood and adult life, the research aims to provide a clearer picture of how "nature" and "nurture" interact in shaping opportunities, with the ultimate goal of informing policies that promote fairer chances for all.
Keywords	Education, genetic endowments, human capital, inequality, intergenerational transmission, polygenic scores, and skill formation
Date received	20 October 2025
Date approved	10 December 2025

Reference	**GDAC_2025_07**
Project title	Decoding Sex-Specific Signatures of Adolescent Aggression Using Deep Learning
Institution	Goethe University Frankfurt (Germany)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Many teenagers with mental health conditions also show serious aggressive behaviours. These behaviours can lead to school exclusion, family stress, and long-term difficulties in adulthood. We want to find out what puts these young people at risk and whether the key factors differ for girls and boys. To do this, we will use data from the Millennium Cohort Study, including answers about family life and friendships plus information on the underlying genetics of the person. We will run computer algorithms that group teens by shared traits and shared risks or experiences. Then we will compare these groups to see which factors (e.g. genetics or peer relationships) matter most for different types of aggression. Our findings could help doctors and families spot warning signs by early adolescence. With clearer risk profiles, they could offer targeted support so that young people can get the right help at the right time.
Keywords	Adolescent aggression, aggressive behaviour, conduct disorder, environmental influences, genetic risk factors, genomic deep learning, and sex-specific analysis

Date received	17 June 2025
Date approved	8 July 2025

Reference	**GDAC_2025_06**
Project title	Genetic and social pathways to adolescent self-esteem
Institution	University of Groningen (The Netherlands)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Adolescent self-esteem is an important concept associated with a range of mental health and social problems, yet its developmental origins remain partly unclear. This project investigates how genetic and social factors jointly influence the development of self-esteem in adolescence. Specifically, it examines how genetic predispositions to externalising problems or lower academic achievement increase the risk of peer victimisation, which, in turn, contributes to lower self-esteem. Using polygenic scores for externalising symptoms and educational achievement, the study will test two separate mediation models to assess both direct and indirect pathways to self-esteem. This approach will provide insight into how genetic and social mechanisms jointly influence self-esteem.
Keywords	Adolescence, gene-environment correlation, self-esteem, and peer victimisation
Date received	17 June 2025
Date approved	8 July 2025

Reference	**GDAC_2025_05**
Project title	Predicting Later Life Outcomes through Longitudinal Cohort Data
Institution	Goethe University Frankfurt (Germany)
Application type	Data access / Genetics
Cohort	NCDS and BCS70
Abstract	This project examines how accurately we can predict individual educational attainment and other life outcomes using data collected at different stages of life. Using British cohort studies, we apply machine learning methods to assess how prediction improves with age. Starting with data available at birth, we then incorporate information from childhood and adolescence.
Keywords	Education, income, and life satisfaction
Date received	16 April 2025
Date approved	8 July 2025

Reference	**GDAC_2025_04**
Project title	Understanding the genetic architecture of grip strength in middle age and older adults
Institution	University College London (UCL)
Application type	Data access / Genetics

Cohort	NCDS and BCS70
Abstract	This project will explore the genetic factors that influence grip strength, which is often used as a measure of overall muscle strength and is an important sign of physical health. Grip strength is closely linked to many different health outcomes. We will conduct a genomewide association study of grip within UK Biobank and seek to replicate findings within the UK birth cohorts (NSHD, 1958BC and BCS70). Importantly, we will not adjust for measures of body size such as BMI, weight or height in our analysis. This will help us better understand the complex relationship between body size and grip strength, and how both are connected to other health factors.
Keywords	Body mass index, genetics, genomewide association study, grip strength, height, and Mendelian Randomisation
Date received	9 April 2025
Date approved	20 May 2025

Reference	**GDAC_2025_03**
Project title	Employing machine learning to better predict the development of mental health problems in young people: a biosocial approach
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	BCS70 and MCS
Abstract	The mental health of young people has been declining over the past two decades. Research has shown that getting early treatment can make symptoms less severe, reduce hospital visits, and lower the chances of symptoms coming back. Previous research has identified a number of early life risk factors associated with mental health problem s, however, using prediction tools for identifying children at risk of mental health problems remains impossible. This project aims to apply machine learning methods to early-life biological (e.g. blood tests, genetics) and social data (e.g. life experiences) in order to create a model that can predict the emergence of mental health problems in young people. By producing prediction models that can accurately identify children at-risk for developing mental health problems in the future, mental health support could be offered in a timely manner to prevent the mental health decline in young people.
Keywords	Machine learning, mental health, prediction, and young people
Date received	27 March 2025
Date approved	9 September 2025

Reference	**GDAC_2025_02**
Project title	Parental income and children's genetic potential for educational attainment
Institution	University of Lausanne (Switzerland)
Application type	Data access / Genetics
Cohort	MCS

Abstract	This study explores how changes in parental income affect children's ability to reach their genetic potential in education. Using data from the Millennium Cohort Study, we focus on a major UK policy change—the 2003 expansion of the Child Tax Credit (CTC)—which provided financial support to lower-income families. By comparing children from families who received additional financial support to those who did not receive any support, we aim to understand whether extra financial resources help children achieve better developmental and educational outcomes. The findings could provide new insights into how economic resources shape child development and long-term success, informing policies that promote equal opportunities for all children, especially those facing financial hardship.
Keywords	Children's education, equality of opportunity, genes, parental income, PGI for education, welfare benefits
Date received	13 March 2025
Date approved	17 April 2025

Reference	**GDAC_2025_01**
Project title	The longitudinal impact of environmental risk factors on depression
Institution	University of Edinburgh
Application type	Data access / Genetics
Cohort	NCDS, BCS70, Next Steps, and MCS
Abstract	To understand depression, it is important to consider change over the life course of individuals as events at different points in life from childhood to old age (such as bullying, divorce, bereavement) influence the likelihood of developing depression. We will test the findings from previous studies using methods which utilise genetic and lifestyle information collected in the same individuals over a long period of time. We aim CLS data access application form 3 to determine which factors influence the onset of depression and tease these apart from factors which frequently coincide with, but do not cause, depression and identify periods of life which are most important for the development of depression and groups of people for which these have the most impact.
Keywords	Causal inference, depression, epigenetics, GWAS, longitudinal study, methylation risk score, and polygenic risk score
Date received	22 January 2025
Date approved	20 May 2025

Reference	**GDAC_2024_13**
Project title	Uncovering Sex-Specific Biological Mechanisms of Depression: Insights From Large-Scale Data Analysis
Institution	King's College London
Application type	Data access / Genetics
Cohort	NCDS, BCS70, and MCS
Abstract	Depression causes low mood and a loss of interest in things that used to be enjoyable. One in six people will suffer from depression during their

	lifetime. However, women are twice as likely to have depression compared to men. There are likely to be a number of reasons for this difference, although we don't yet fully understand them. Genes are likely to play a role in this difference between the sexes and my research using NCDS, BCS70, and MCS data will explore this area in more detail by examining the sexes separately. These results will be combined with those from other studies to provide the strongest clues yet about depression genetics. We will examine the genome for areas that differ between the sexes for their role in depression potentially leading to better individualised treatments.
Keywords	Depression, genes, genetics, mental health, and sex differences
Date received	24 September 2024
Date approved	4 November 2024

Reference	**GDAC_2024_12**
Project title	How genetic liability affects the development of mental health symptoms
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Many factors may come together to increase one's risk of developing a mental illness, including genes and biological markers, as well as aspects of the social environment. We know that many mental illnesses run in families, and that genetics can impact brain development and mental health symptoms. However, not all individuals who are at high risk genetically will go on to develop a mental illness. It is therefore important that we try to better understand how a person's genetic risk can be shaped by social or environmental factors. Through identifying ways that genetic risk is impacted by broader factors may create opportunities to prevent the onset of serious mental illnesses. The goal of this research is to better understand how genes influence the development of serious mental illnesses over time, including possible mediators which might affect susceptibility.
Keywords	Gene-environment interaction, genetics, and mental illness
Date received	24 September 2024
Date approved	22 October 2024

Reference	**GDAC_2024_11**
Project title	Bio-social pathways to poor mental health: using blood samples to index neurobiological factors
Institution	Institute for Social and Economic Research (ISER)
Application type	Data access / Genetics
Cohort	NCDS, BCS70, and MCS
Abstract	Genetic risk for poor mental health can be measured in different ways. A common approach is to create a polygenic index (PGI) by adding up a person's unique set of genetic variants which have been shown to

	increase risk for a mental health disorder (like depression, schizophrenia, etc). One problem with disorder-PGIs is that it's not always clear which biological mechanisms they represent. Another problem is that the environments we grow up in – which might also increase risk for poor mental health – are often correlated with disorder-PGIs. In other words, disorder-PGIs measure risk for poor mental health, but the mechanisms aren't always clear. This project will instead make "endophenotype"-PGIs: PGIs which measure different aspects of brain structure and function, i.e. plausible biological mechanisms affecting our mental health. We will compare the ways in which disorder-PGIs and endophenotype-PGIs are associated with both mental health and the environments people grow up in.
Keywords	Endophenotypes, genetics, mental health, neuroscience, and socio-genomics
Date received	3 July 2024
Date approved	26 September 2024

Reference	**GDAC_2024_10**
Project title	Understanding the combined impact of behavioural and neurodevelopmental disorders on learning and child development
Institution	Queen Mary, University of London
Application type	Data access / Genetics
Cohort	MCS
Abstract	The combined impact of neurodevelopmental and behavioural disorders on learning and child development is critically under-researched. Our previous research has shown that a high proportion of genetic and environmental risk is shared across neurodevelopmental and behavioural disorders. Yet, there is a major gap in our understanding of their co-occurrence, which has largely only been investigated between ADHD, autism, and conduct disorder. With this project we will: 1) Map the co-occurrence between all neurodevelopmental and behavioural disorders in youth. 2) Quantify their joint impact on learning outcomes, unravelling genetic and environmental pathways. 3) Investigate how children with a higher transdiagnostic genetic risk might experience, evoke and select environments that relate to educational outcomes and pinpoint protective factors that foster learning.
Keywords	Academic achievement, behavioural problems, education, gene-environment correlation, neurodevelopmental disorders, and transdiagnostic risk
Date received	30 June 2024
Date approved	6 August 2024

Reference	**GDAC_2024_09**
Project title	Towards an understanding of social media: The effects on health, well-being, and genetic susceptibilities
Institution	NOVA University Lisbon (Portugal)

Application type	Data access / Genetics
Cohort	MCS
Abstract	In 2022, 63% of the world population actively used social media, spending a daily average of 2 hours and 27 minutes on these platforms. Despite its widespread adoption, our understanding of social media's impact on health, well-being, and human capital remains limited, mostly due to lack of data availability on social media use. This proposal has two objectives. The first one proposes to estimate the causal impact of social media on mental health and well-being by exploiting the timing of exogenous shocks that influence social media exposure. The second investigates whether genetic susceptibilities towards mental disorders and educational attainment explain patterns of social media use. The exploration of these objectives will contribute towards advancing our comprehension of social media.
Keywords	Causal effects, genetic susceptibilities, health, social media, well-being, and within-family analysis
Date received	11 June 2024
Date approved	26 June 2024

Reference	**GDAC_2024_08**
Project title	Intergenerational educational mobility: a sociogenomic approach
Institution	University of Groningen (Netherlands)
Application type	Data access / Genetics
Cohort	NCDS and MCS
Abstract	To what extent educational outcomes are transmitted across generations is a key indicator of equality of opportunity in a society, but little is known about the roles played by genes and parental environment in this transmission. In this project we aim to disentangle the relative role of genetic endowments and parental investments in predicting offspring's educational attainment, using data from two British birth cohort studies (NCDS and MCS). This will allow us to assess whether and how this gene-environment interplay has changed over time. Moreover, by looking at offspring's genetic endowments for distinct dimensions (cognitive, non-cognitive, and neurodevelopmental traits) that are predictive of educational attainment, we will assess what dimension (if any) is more relevant for educational attainment and how the role of each dimension changes depending on parental education and parental investment.
Keywords	Educational attainment, educational mobility, gene-environment correlation, gene-environment interaction, non-cognitive skills, and parental investments
Date received	5 June 2024
Date approved	20 June 2024

Reference	**GDAC_2024_07**
Project title	Genetic and labour market influences on health over the life cycle

Institution	University of Bologna (Italy)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Health inequality begins at conception and amplifies throughout the life cycle. Disparities in mental and physical health stem from a blend of genetic predispositions and environmental influences. Genetics may either mitigate or exacerbate the effects of the environment. This project integrates methodologies from genetics and economics to examine how gene-environment interactions (GxE) shape health outcomes across individuals' lifespans. Our focus will be on analysing the ramifications of exposure to challenging labour market conditions, which will be measured by automation, the process in which industrial robots replace human tasks in production. By studying exposure to these environments at various life stages, we aim to capture critical periods that influence health outcomes and determine whether the GxE effect persists over the short or long term.
Keywords	Causal inference, foetal origin hypothesis, and social genomics
Date received	3 June 2024
Date approved	6 August 2024

Reference	**GDAC_2024_06**
Project title	Mapping the longitudinal effects of gene variants linked to psychosis and schizophrenia
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	NCDS, BCS70, and MCS
Abstract	Individuals with psychosis and schizophrenia face many challenges throughout their lives. Recent research has shown that certain changes to DNA (which we call, mutations) can increase the likelihood of a person developing these conditions. Currently, we do not know whether these mutations affect other aspects of a person's life. In this project, we hope to explore this. We will study individuals who have inherited these mutations, and we will assess their physical, mental, and cognitive health at sequential timepoints throughout childhood, early adulthood, and in some instances, through to late middle age. This will help us to map the long-term consequences, if they exist, of living with these mutations. Ultimately, through this work, we hope to identify new ways to support these individuals so that they can live positive, fulfilling, and healthy lives.
Keywords	Cognitive functioning, disease risk, gene-environment interactions, genetic epidemiology, psychosis, and schizophrenia
Date received	3 May 2024
Date approved	3 June 2024

Reference	**GDAC_2024_05_gwas**

Project title	Improving use of the CLS cohorts in Genomewide Association Studies (GWAS)
Institution	University College London (UCL) – CLS
Application type	Data access / Genetics
Cohort	NCDS, BCS70, and MCS
Abstract	Large scale Genomewide Association Studies (GWAS) are the mainstay of modern genetic knowledge, helping to identify the points of variation in the human genome that are related to traits ranging from blood pressure to BMI to personality. GWAS require prediction samples to ensure that their results replicate and do not just reflect factors related to their discovery samples. This project will enable the CLS cohorts to be more efficiently used in GWAS, offering nationally representative samples with high response rates for holdout/prediction.
Keywords	Collaboration, genetics, GWAS, health, and social
Date received	10 April 2024
Date approved	10 April 2024

Reference	**GDAC_2024_04**
Project title	Adverse childhood experiences and adolescent mental health: A propensity score analysis
Institution	University of Oxford
Application type	Data access / Genetics
Cohort	MCS
Abstract	Numerous studies have demonstrated strong associations between adverse childhood experiences (ACEs) and mental health problems during adolescence. However, despite the well-established links between ACEs and mental health problems, less is understood regarding their causal relationship. Previous research which mainly relied on epidemiological methods faced limitations such as genetic and environmental confounding. For example, children exposed to ACEs might already have genetic predispositions for mental health problems, or they might be experiencing other environmental risks simultaneously. This project will use causal inference methods to control for these confounding factors, and better understand the causal relationship between ACEs and mental health problems. It is crucial to understand the causal relationship because 1) if causal effects between certain ACEs and mental health problems are found, this could inform interventions to target the specific adversities and prevent children from developing mental health problems; 2) if non-causal effects are found, this could inform interventions to focus on treating pre-existing vulnerabilities among exposed children.
Keywords	Adolescence, adverse childhood experiences, causal inference, mental health, and propensity score analysis
Date received	4 April 2024
Date approved	11 July 2024

Reference	GDAC_2024_03
Project title	Early conditions, delayed adult effects, and morbidity, disability, and mortality in modern human populations (ECHO)
Institution	University of Wisconsin-Madison (United States)
Application type	Data access / Genetics
Cohort	NCDS, BCS70, and MCS
Abstract	The subproject (within the larger ECHO project) aims at assessing the influence that parental household and life course characteristics exert on individuals' BMI, obesity, and Type 2 Diabetes. We propose to do this by using (i) individuals' fixed genotypic features (Polygenic Risk Scores); (ii) early conditions (birthweight, maternal health status, parental household conditions), (iii) individuals' traits that vary over the life course (including marital status, educational attainment, labour force). We aim to separate as neatly as possible effects associated with individual's genotypes, early life, and parent-of-origin influences, and modifying factors that unfold over the lifetime of individuals. Three key questions: (i) direct genetic effects and interaction effects with early conditions (low birth weight, parental characteristics? Do these change over time); (ii) BMI and obesity trajectories over a life course: What is in the 'gross' (including genetic predisposition) and ('net' (excluding genetic predisposition) intragenerational correlation of BMI and obesity, and (iii) do different cohorts (1958, 1970, 2000) differ?
Keywords	Early conditions, environment interactions, indirect genetic and niche effects, gene-environment interactions, obesity, and type 2 diabetes
Date received	21 March 2024
Date approved	20 June 2024

Reference	GDAC_2024_02
Project title	Sociogenomics, gene-environment interplay, education, inequalities, life course, social mobility, intergenerational transmission.
Institution	University of Oxford
Application type	Data access / Genetics
Cohort	NCDS, BCS70, and MCS
Abstract	This study aims to understand to what extend inequalities in education attainment are explained by variations in genetic endowments, social origins, and the interaction between both. Its objective is to unpack educational careers by exploring at which stages genetics and social origins matter most for low- and high-socioeconomic status students, and how this might have changed across birth cohorts. Incorporating genetic data to solve this puzzle has methodological and conceptual implications – it will improve existing ability measures by constructing polygenic scores and will allow to revisit long-lasting hypothesis in social research. The core method will consist of a series of regression analysis on the completion of key academic transitions, accounting for genetic, family, and school-level covariates. By abandoning rigid 'nature vs nurture' divides, this research will provide a comprehensive

	understanding of education inequalities, in addition to providing insights into how policy interventions can mitigate the influence negative of social environments.
Keywords	Education, gene-environment interplay, inequalities, intergenerational transmission, life course, sociogenomics, and social mobility
Date received	1 February 2024
Date approved	20 March 2024

Reference	**GDAC_2024_01**
Project title	GWAS of political phenotypes
Institution	University of Minnesota, Twin Cities (United States)
Application type	Data access / Genetics
Cohort	NCDS, BCS70, and MCS
Abstract	Political scientists have long sought to comprehend the myriad factors that drive citizens to participate in elections and choose their favoured candidates. While socio-economic and environmental factors play a significant role, twin studies have suggested that much of the variation in political phenotypes may be accounted for by genetic variation. However, why genetic variation is associated with political traits and which genes matter is unclear. In other words, the etiology, psychological and socio-economic causes for political traits are uncertain. We intend to illuminate the etiology of political traits with genome-wide association study of voter turnout and vote choice to identify genetic markers. The aim of this research project is to unravel the genetic underpinnings of voter turnout, a complex behavior influenced by a combination of environmental, social, and genetic factors. Understanding the genetic basis of political participation has the potential to contribute valuable insights into the mechanisms shaping civic engagement.
Keywords	GWAS, politics, polygenic scores, and voting
Date received	29 January 2024
Date approved	20 June 2024

Reference	**GDAC_2023_22**
Project title	Unpacking the Emergence of Dementia Etiology Across the Life Course
Institution	Purdue University (United States)
Application type	Data access / Genetics
Cohort	NCDS and BCS70
Abstract	We will analyse the links between genotypes and phenotypes to investigate the onset times of risk factors for Alzheimer's and Dementia (AD) across the life course using data from the CLS. Our insights will focus on pinpointing specific periods when these outcomes manifest as individuals age. We hypothesise that genetic data and structural equation modelling can help identify the specific times when Alzheimer's risk factors emerge as individuals age throughout their lives.

Keywords	Alzheimer's Disease Related Dementias, Genetics, Genome-Wide Association Study, Genomic Life Course Model, and Genomic Structural Equation Modelling
Date received	29 December 2023
Date approved	16 July 2024

Reference	**GDAC_2023_21**
Project title	Gene-environment interplay on the development of language and reading-related skills and co-occurring traits
Institution	University of Edinburgh
Application type	Data access / Genetics
Cohort	BCS70
Abstract	Being able to read, write and speak clearly is related to life success. How well people can read/write is partly due to their genes. Because reading is a learned skill genes will interact with a person's environment. In the 1970s British Cohort Study, we will test whether a person's genes are linked to their reading, speech, and language abilities, and to co-occurring conditions like developmental co-ordination and numeracy difficulties. By combining these genome-wide association results with those from the international Genetics of Language and other Consortia we will help to discover new genes associated with reading, speech, language, and co-ordination problems. We will further test genetic risk for dyslexia alongside measures like parental book reading, occupation and income. This will help us understand whether there are aspects of the home and school environment that can lessen the genetic influence on reading difficulties. We will check whether these are the same across childhood and youth, how they relate to co-occurring traits, like ADHD, and if they are related to health outcomes.
Keywords	Genetics, language, literacy environment, reading skill, and speech problems
Date received	21 December 2023
Date approved	21 May 2024

Reference	**GDAC_2023_20**
Project title	Internal access: derived genetic and epigenetic variables for CLS cohorts
Institution	University College London (UCL) – CLS
Application type	Data access / Genetics
Cohort	NCDS, BCS70, Next Steps, and MCS
Abstract	Creating derived summary variables from complex genetic data requires specialist knowledge and large-scale computing resources. However, these derived summary variables are easy to use and can be enormously useful to researchers, enabling non-specialists to use genetically informed research designs. For example, polygenic scores provide anonymous measures of genetic predisposition to specific traits which summarise information contained across many hundreds or

	thousands of genetic variants. Similarly, epigenetic clocks provide information of 'biological aging' summarised from epigenetic data. Derived genetic and epigenetic variables therefore inform an important part of CLS' strategy and enable an agreed output of our current ESRC funding. This project is proposed to formally approve a pathway for CLS researchers to access derived genetic data internally. The overarching goal is to help build CLS' genetics resource for the wider benefit of the research community.
Keywords	Cohort resources, derived variables, epigenetic clocks, and polygenic scores
Date received	28 November 2023
Date approved	6 December 2023

Reference	**GDAC_2023_19**
Project title	Examining the association between emotional and neurodevelopmental problems in young people and adult health outcomes; the impact of heterogeneity and genetic factors
Institution	Cardiff University
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Childhood emotional problems are associated with mental ill health, cardiovascular risk, and social functioning in adulthood. This project will extend understanding by tracking outcomes to age 50 and by testing the role of factors that explain why some children with emotional problems experience poor outcomes and others not. The proposal has three aims: 1) examine associations between child/adolescent emotional problems and adult cardiovascular risk (mid-life) and mortality, adult mental health (Malaise trajectories, 23-50 years), and social outcomes (up to 50 years); 2) test the role of age at onset and persistence of emotional problems, common and rare genetic risk factors, childhood social adversity, and co-occurring childhood symptoms (ADHD, irritability); 3) examine the role of genetic and other childhood risks in explaining links between adult cardiovascular risk and mental health.
Keywords	ADHD, emotional, cardiovascular, genetic, and mental health
Date received	23 November 2023
Date approved	8 January 2024

Reference	**GDAC_2023_18**
Project title	Impulse Encoded: Using Mendelian Randomization to Identify the Causal Impact of Time and Risk Preferences on Health Behaviors and Healthcare Costs
Institution	Imperial College London
Application type	Data access / Genetics
Cohort	BCS70
Abstract	Our research seeks to answer the following question: does the way individuals value immediate costs or benefits relative to future costs or

	benefits, known as 'time preference', and their willingness to take risks, also known as 'risk tolerance', have an impact on their behaviours and choices? We are especially interested in behaviours and choices that may lead to poor health in the future. We are also seeking to understand whether time and risk preferences may explain why people of lower socio-economic status are more likely to engage in unhealthy behaviours and incur higher healthcare costs. To investigate these questions, we will analyse data from the British Cohort Study 1970, which includes genetic information, health administrative records, and lifestyle details. Previous studies have shown that people who have certain genetic characteristics are more likely to place greater importance on immediate costs and benefits and exhibit a higher propensity for risk-taking behaviours. Since these genetic characteristics are distributed in a random way in the population, we plan to use information on those characteristics to compare behaviours in those who have and those who do not have those characteristics, to understand the role of time and risk preferences in shaping such behaviours. If we can show that time and risk preferences have a role, we will assess whether they are responsible for differences in behaviour and financial burdens among people who belong to different social classes.
Keywords	Health behaviours, health cost, health inequalities, Mendelian Randomisation, risk preference, socioeconomic status, and time preference
Date received	22 November 2023
Date approved	26 March 2024

Reference	**GDAC_2023_17**
Project title	What is the Source of the Health Gradient? The Case of Obesity
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	BCS70 and MCS
Abstract	Why are those who grew up in poorer families more likely to become obese than those from richer backgrounds? This project aims at understanding how health behaviours and genetic predispositions interact from childhood until mid-adulthood in generating health differences. We will estimate a semi-structural dynamic model which models how obesity develops via early differences in childhood health investments and endowments, differences in education attainment and wages, and differences in individuals' own health behaviours in adulthood. Genes play a potentially important role, both in driving differences in the health process, but also by driving differences in investments; the interactions of investments and genetic endowments may be a crucial aspect in
Keywords	Health investments, human capital development, lifecycle model, and obesity
Date received	22 November 2023
Date approved	31 January 2024

Reference	GDAC_2023_16
Project title	Parental investment responses to child endowments
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	MCS
Abstract	In this research project, we investigate how parents adjust their investments in children and the environment in which their children grow up in response to the genetic endowment of their offspring and their own. Thereby, we pay particular attention to changes over the child's lifecycle, differential investment patterns by mothers and fathers, and interactions with environmental variables such as the quality of neighbourhoods and schools. As a result of this project, we will be able to characterize whether parents in the UK reinforce or compensate the genetic endowments of children, when such reinforcing/compensating behaviour is likely to emerge (differences by child age), which parent is particularly likely to do so (differences between mothers and fathers), and under which environmental conditions such reinforcing/compensating behaviour is particularly pronounced (differences by quality of neighbourhoods and schools).
Keywords	Gene-by-environment interplay, parental investments, polygenic indices, skill formation, and social science genetics
Date received	15 November 2023
Date approved	3 January 2024

Reference	GDAC_2023_15
Project title	Influence of Genetic and Environmental Factors on the Development of Individuals Born Preterm or Low Birth Weight
Institution	University of Warwick
Application type	Data access / Genetics
Cohort	MCS
Abstract	Around 1 in 10 babies worldwide are born prematurely, before 37 weeks of pregnancy. Premature birth or with low birth weight (< 2.5 kg) can result in long-term challenges, affecting physical and mental health, as well as social functioning. However, some individuals exhibit better-than-expected outcomes. Our research seeks to address the vital question of how genetic and environment (such as family, school, social, and community) interact in shaping the development of these individuals. The Millennium Cohort Study data is invaluable in unravelling these complex connections. By analysing genetic data and accessible survey data, we aim to identify genetic markers and protective factors that associated with resilience in mental health and social functioning outcomes. This investigation may pave the way for more precise and personalized healthcare strategies and promote a strengths-based approach to understanding and supporting the well-being of this population.

Keywords	Environment, genetics, low birth weight, preterm birth, resilience, and risk
Date received	15 November 2023
Date approved	14 December 2023

Reference	**GDAC_2023_14**
Project title	Gene-environment interplay on the development of language and reading-related skills and co-occurring traits
Institution	University of Edinburgh
Application type	Data access / Genetics
Cohort	MCS
Abstract	Being able to read, write and speak clearly is related to life success. How well people can read/write is partly due to their genes. Because reading is a learned skill genes will interact with a person's environment. In the Millennium Cohort Study, we will test whether a person's genes are linked to their reading, speech, and language abilities, and to co-occurring conditions like developmental co-ordination and numeracy difficulties. By combining these genome-wide association results with those from the international Genetics of Language and other Consortia we will help to discover new genes associated with reading, speech, language, and co-ordination problems. We will further test genetic risk for dyslexia alongside measures like how a person is taught to read, and parental book reading, occupation and income. This will help us understand whether there are aspects of the home and school environment that can lessen the genetic influence on reading difficulties. We will check whether these are the same across childhood and youth, how they relate to co-occurring traits, like ADHD, and if they are related to health outcomes.
Keywords	Genetics, language, literacy environment, reading skill, and speech problems
Date received	15 November 2023
Date approved	6 December 2023

Reference	**GDAC_2023_13**
Project title	The protective effect of social support against mental health difficulties for adolescents with and without experience of adversity, a genetically sensitive approach
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	This study will look at the relationship between social support at age 14 and mental health and wellbeing outcomes at age 17, using data from the Millennium Cohort Study. Firstly, we will explore the relationship between social support and mental health for all adolescents, before exploring whether effects are different for adolescents who have experienced childhood adversity (e.g., physical abuse and neglect). We will use genetic data to consider whether social support is predicted by

	genetic factors and whether associations between social support and mental health outcomes are explained by genetics.
Keywords	Adolescence, genetically informed, mental health, protective factors, social support, and wellbeing
Date received	25 October 2023
Date approved	21 December 2023

Reference	**GDAC_2023_12**
Project title	How do parental investments react to child endowments? Evidence from the MCS
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	How do parents allocate resources across their children? This is a longstanding question in social science. Faced with their children's characteristics, their budgets, their preferences, and their beliefs, parents can behave in a number of different ways, affecting cross sibling inequality. For example, they may decide to invest more in the most able children, they may decide to compensate lower endowments with stronger investments, they may decide to equalize investments or outcomes, and so on. In this paper we will investigate the extent to which parents react to the genetic endowments of their children, taking advantage of the fact that, conditional on parents' genetic endowments, children's endowments are random. We will also examine how these behaviours change as children age and parents observe more about their children.
Keywords	Child development, genetics, and parental investments
Date received	25 October 2023
Date approved	15 February 2024

Reference	**GDAC_2023_11**
Project title	Exploring antecedents of academic achievement in the MCS
Institution	University College London (UCL) – CLS
Application type	Data access / Genetics
Cohort	MCS
Abstract	The factors that influence how well children perform throughout schooling are extremely complex. There is evidence that experience of the school environment, such as school enjoyment, is strongly related to school performance and independent of family socioeconomic background. There is also evidence that genetic factors play a role in school performance, but that these operate in a complex manner and can operate through the family environment. We will use the data in this project to answer the following questions: 1. How does school enjoyment relate to academic achievement in the MCS? 2. Does school enjoyment have a genetic component?

	3. How do genetic and environmental associations with school performance operate?
Keywords	Education, genetics, longitudinal, MCS, and school enjoyment
Date received	20 October 2023
Date approved	10 January 2024

Reference	**GDAC_2023_10**
Project title	Finding the "missing environmentality"
Institution	University College London (UCL) – CLS
Application type	Data access / Genetics
Cohort	NCDS, BCS70, and MCS
Abstract	Can we explain human behaviour? Social science theories are often focused on specific factors influencing our life outcomes such as parental education shaping children's educational success. However, recently, researchers started to ask bigger questions. Overall, to what extent can we explain or predict outcomes such as education, well-being, or family planning in our life courses with all our data? Is life deterministic or does chance shape our destinies? Is our biology or our social environment more important and how do they play together? Answering such questions is not only important for social science theory building and our idea of how stable the social world is, it also tells us how and to what extent we can intervene to improve our life qualities and health. This project infuses social sciences with knowledge from genetics and in particular their complex statistical methodology to answer old and persisting questions in novel ways. We will be able to disentangle genetic from social factors and take gene-environment interaction into account and quantitatively answer questions such as: Can we explain human behaviour?
Keywords	Education, fertility, missing heritability, social prediction, and well-being
Date received	26 September 2023
Date approved	16 July 2024

Reference	**GDAC_2023_09**
Project title	Investigating environmental mediation of genetic effects in the MCS
Institution	University College London (UCL) – CLS
Application type	Data access / Genetics
Cohort	MCS
Abstract	Children inherit not only their genomes but also their environments from their parents. This can lead to spurious correlations between genes and environments that biases research, but it can also be leveraged to better understand how the environments that parents create affect their children.
Keywords	Education, environment, families, genetics, and mediation
Date received	22 September 2023
Date approved	10 January 2024

Reference	GDAC_2023_08
Project title	Multi-ancestral GWAS meta-analysis of childhood growth and overweight (EGG consortium)
Institution	University of South Australia (Australia)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Birth weight, patterns of growth and childhood obesity are important determinants of health with effects extending into adulthood. A better understanding of the factors driving early growth phenotypes, such as birth weight, patterns of growth and childhood obesity will help to identify intervention strategies, thereby potentially improving health outcomes across the lifespan. Several studies have shown that growth and predisposition to obesity are heritable traits, and for example, genome-wide meta-analyses published by the Early Growth Genetics (EGG) consortium has led the discovery of genetic loci regulating birth weight, gestational age, growth and obesity during the childhood and puberty. The EGG consortium is now updating meta-analyses of childhood obesity and overweight and this application is to allow the incorporation of 1958 British Birth Cohort data to these analyses, where it will form an integral part of the overall EGG meta-analysis.
Keywords	Birth weight, BMI, childhood, copy number variation, growth, GWAS, multi-ancestry, obesity, and overweight
Date received	6 September 2023
Date approved	11 October 2023

Reference	GDAC_2023_06
Project title	Changes in the Spatial Distribution of Genes Linked to Education and Cognitive Ability: A Cross-Cohort and Life Course Analysis of Four British Birth Cohort Studies
Institution	University College London (UCL) – CLS
Application type	Data access / Genetics
Cohort	NCDS, BCS70, and MCS
Abstract	Regional disparities in economic output have increased across Britain since the 1980s and are among the most unequal in the developed world. One reason for these disparities may be differences in human capital. Cognitive ability and education are strong predictors of job performance and multiple measures of economic productivity. Moreover, these are highly heritable and polygenic traits. Changes in the spatial distribution of genetic liability to these traits (as measured with polygenic scores) may have contributed to growing disparities in economic performance, e.g., via selective migration. This project will use data from four British birth cohort studies (born in 1946, 1958, 1970, and 2001/2002, respectively) to examine how the distribution of polygenic scores has changed across cohorts and over participants' life courses. We will measure spatial distribution directly using postcode-level data,

	and indirectly, by linking postcodes to area-level measures related to economic production, including deprivation and unemployment rates.
Keywords	Cognitive ability, education, genetics, migration, and selection effects
Date received	5 June 2023
Date approved	17 October 2023

Reference	**GDAC_2023_05**
Project title	Understanding factors related to age of diagnosis of neurodevelopmental conditions
Institution	University of Cambridge
Application type	Data access / Genetics
Cohort	MCS
Abstract	We want to understand why some people receive a diagnosis of neurodevelopmental conditions (e.g., Autism and ADHD) during childhood and others during adolescence, and what the consequence of late diagnoses are on mental and physical health. To achieve this, we will investigate social, developmental, and genetic factors that are linked to age of diagnosis and trajectories of neurodevelopmental traits.
Keywords	ADHD, autism, diagnosis age, neurodevelopment, and SDQ
Date received	11 April 2023
Date approved	21 July 2023

Reference	**GDAC_2023_04**
Project title	Intergenerational predictors of language, communication and social abilities in children and adolescents and their mediating effect on later-life health and development - a genetic investigation
Institution	Max Planck Institute for Psycholinguistics (Netherlands)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Language is a distinct and complex human feature that plays a crucial role during a child's development and has been linked to later behavioural, cognitive, social and health outcomes. However, the role of parental genes that directly or indirectly mediate this relationship, especially within typically developing children, is still not well understood. This project aims to identify mechanisms across parent and child generations contributing to the development of children's language, communication and social abilities using genetic tools. In turn, we aim to use identified genetic information from the parent generation as robust, confounding-free instruments to study overarching relationships with children's language, communication and social abilities and later life outcomes. Thus, this research can help predicting generational trajectories shaping future health and development in children and adolescents and inform educational tasks and health interventions aiming to improve later-life health and wellbeing in children.

Keywords	Direct and indirect genetic effects, education, genetic association, health, home environment, language, lifestyle, literacy, parent child transmission, parenting, social behaviour, social communication, and social interaction
Date received	28 February 2023
Date approved	9 May 2023

Reference	**GDAC_2023_03**
Project title	Examining the effect of autistic trait trajectories on social, health and wellbeing outcomes across multiple cohorts with genetically informed methods
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	MCS
Abstract	Little is known about how social communication difficulties in the general population change over time and how these changes impact health, wellbeing and social outcomes. If we can better understand how social communication difficulties change over time this may help us to find age ranges where interventions may have the most positive impact on these outcomes. We will look at how social communication difficulties change over time, using data from five existing large studies and we will combine this information across studies to give us a better picture of changes in the general population. Then, using the same studies, we will use genetic information and the information about changes in social communication difficulties over time in genetically informed methods to see how changes in social communication difficulties over time impact our outcomes of interest. We will also look at these relationships using observational analyses.
Keywords	Autistic traits, trajectories, health outcomes, social outcomes, and wellbeing
Date received	22 February 2023
Date approved	20 April 2023

Reference	**GDAC_2023_02**
Project title	Exploring population phenomena in genetic associations
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	MCS
Abstract	The aim of this project is to use genetic data from the Millennium Cohort Study to explore how population phenomena can bias estimates from genetic studies, making them less reliable. Population phenomena include processes like assortative mating, where individuals select mates with similar characteristics instead of at random. We will use Genome-wide Complex Trait Analysis (GCTA), which estimates genetic heritability of complex traits, such as education and socioeconomic position. This works by comparing how similar are unrelated individuals

	genetically, and how similar they are in terms of complex traits. We can infer that the measured genetics are likely to influence a particular trait when unrelated individuals who are similar genetically also have similar trait measurements.
Keywords	Education, GCTA, genetic association studies, population phenomena, and sociological factors
Date received	13 January 2023
Date approved	20 April 2023

Reference	**GDAC_2023_01**
Project title	Gene Environment Interactions in Mental Health Trajectories of Youth
Institution	Maastricht University (Netherlands)
Application type	Data access / Genetics
Cohort	MCS
Abstract	We have growing evidence for the involvement of genes, several environmental factors, and their interactions in mental health outcomes. Now, we need to puzzle out the specific psychological, developmental, biological, and genetic mechanisms that can help us understand how these mechanisms make some people more susceptible and others more resilient. In this project, we aim to answer these questions by estimating and investigating trajectories of mental health in young people. We will first model dynamic mental health trajectories by using traditional statistical methods, as well as machine learning. We will then investigate how environment and genes influence these mental health trajectories. We will apply novel biologically-informed and symptom-specific gene sets to better understand specific mechanisms of gene-environment interaction. We will also study the dynamic, interactive, and time-dependent influence of a range of "non-genetic" environmental factors, such as childhood adversities, stressful life events, peer-bullying, and drug use, on mental health outcomes.
Keywords	Exposome, gene-environment interaction, mental health, trajectory, and trans-syndromes
Date received	13 December 2022
Date approved	15 June 2023

Reference	**GDAC_2022_23**
Project title	Environmental impact on the relationship between ADHD genetics and physical health conditions in childhood
Institution	University of Southampton
Application type	Data access / Genetics
Cohort	MCS
Abstract	Research has shown that people with ADHD are more likely to have other health conditions, such as asthma and epilepsy. These relationships are complex and may be influenced by several

	environmental risk factors. Many conditions, including ADHD, have a genetic origin and are known to be passed on from parent to child. Sometimes, a parent may have a heritable condition that their child does not inherit but which still influences the child's environment growing up. This indirect effect of such non-transmitted genetic risk is known as genetic nurture. This project will explore ADHD risk via genetic nurture and direct genetic transmission using information from parents and children. We aim to examine the association between ADHD and physical health, and the environmental factors that may impact this relationship.
Keywords	ADHD, polygenic risk scores, physical health, environmental factors, genetic risk, and genetic nurture.
Date received	14 December 2022
Date approved	26 January 2023

Reference	**GDAC_2022_22**
Project title	GxE interaction: Decomposing the polygenic score for education into its component
Institution	ENSAE Paris (France)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	This project addresses the question of whether and how family's socio-economic status (SES) interacts with the genetic predisposition for education in shaping educational outcomes. The aim is to disaggregate the polygenic index (PGI) for education into its' components. Indeed, the PGI for education may predict at the same time cognitive performances but also depression or other disorders (i.e., bipolar disorder). We want to find out if advantaged families boost the part of PGI for education that predicts cognitive performances (hypothesis 1) and compensate for the part that predicts depression or other disorders (hypothesis 2). For this, we first simulate different scenarios, and then using NCDS we test our hypotheses.
Keywords	Genes-environment interaction (GxE), polygenic index (PGI) for education, decomposition, educational inequalities, compensatory advantaged hypothesis, and Scarr-Rowe hypothesis
Date received	20 November 2022
Date approved	8 December 2022

Reference	**GDAC_2022_21**
Project title	Genomic and environmental influences on mental health risk and resilience
Institution	Massachusetts General Hospital (United States)
Application type	Data access / Genetics
Cohort	MCS

Abstract	We plan to use data-driven methods to examine how wide-ranging environmental and genetic factors are associated with youth mental health outcomes (risk and resilience) in the Millennium Cohort Study.
Keywords	Depression, exposome, genome, resilience, and youth mental health
Date received	16 November 2022
Date approved	12 May 2023

Reference	**GDAC_2022_20**
Project title	Genome-wide association study (GWAS) of Personality
Institution	University College London (UCL) – CLS
Application type	Data access / Genetics
Cohort	NCDS and MCS
Abstract	Our aim is to contribute the CLS data to several genetic consortia. Genetic consortia have been developed to use existing genetic data and harness the power of sample size in order to identify instances where certain traits have a genetic cause. We aim to contribute to the Revived Genetics (of) Personality Consortium (RGPC) by investigating if genetic variation is associated with personality. We will be utilising the Big Five taxonomy: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness. We will be using genetic data from MCS and NCDS in combination with questionnaire answers relating to personality.
Keywords	GWAS, Personality, and the Big Five
Date received	2 November 2022
Date approved	10 November 2022

Reference	**GDAC_2022_19**
Project title	Identifying common maternal variants related to the occurrence of Cleft in their children
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	MCS
Abstract	We have performed one of the first studies of maternal genetic variation linked to cleft lip and or palate in children using genetic data from 2500 mothers whose children were born with a cleft and comparing them to mothers from the Millennium Cohort Study. In addition, we have also performed a GWAS of cleft cases and controls using genetic data from Cleft Collective participants and children from the Millennium Cohort Study as controls. We are now requesting to use the MCS genetics data already in Bristol to perform follow-up analysis. Our aim is to test if children born with cleft and their mothers are more or less likely to carry common genetic variation linked to neurodevelopmental and mental health problems. We have previously identified that children born with cleft lip and or palate from the Cleft Collective are more likely to have early behavioural problems (Berman et al. 2022) and in this project we

	will investigate if this higher risk could be due to shared genetic variation with mental health and neurodevelopmental disorders.
Keywords	Genetic risk factors, GWAS, novel loci, orofacial cleft, and mothers
Date received	1 November 2022
Date approved	25 January 2023

Reference	**GDAC_2022_18**
Project title	Genetic Mediation of Predictors of Executive Function Development
Institution	University of Edinburgh
Application type	Data access / Genetics
Cohort	MCS
Abstract	Executive function refers to cognitive skills such as impulse control and updating of working memory. Multiple studies support a relation between executive function deficits during development and maternal smoking, low socio-economic status, and single-parent status. However, twin studies of executive function indicate that differences are almost entirely due to genetics, with only minor contributions from environmental factors. These two branches of research thus disagree as to the relative contributions of genetics and environment. This project will use molecular genetic data from individuals and their parents to unravel this relationship. By comparing genetic predictors of environmental factors with those of executive function, it will be possible to determine if shared genetics explain their correlations. Parental genetics will then be used to determine if this is best explained by shared genetics between parent and child shaping outcomes for both, or by parental genetics shaping the environment for their children.
Keywords	Executive function, behavioural genetics, genetic nurture, maternal smoking, and single-parent households
Date received	27 October 2022
Date approved	26 January 2023

Reference	**GDAC_2022_17**
Project title	Early mediators of genetic influences on educational attainment
Institution	University of Cambridge
Application type	Data access / Genetics
Cohort	MCS
Abstract	This project will investigate two research questions: (1) What are the early manifestations of genetic influences on educational abilities and externalising problems? (2) Do children's education-associated and externalising-associated genes influence the early caregiving they receive, in turn influencing their later externalising and educational outcomes? The first research question will be addressed by examining whether (a) the effects of children's education polygenic scores on their educational performance are mediated via their early language and cognitive skills and internalising/externalising problems; (b) the effects of children's externalising polygenic scores are mediated via their early

	emotional dysregulation. The second question will be addressed by examining whether (a) the effects of children's education polygenic scores on their educational performance are mediated via the home learning environment; (b) the effects of children's externalising polygenic scores on their externalising problems are mediated via parental discipline practices.
Keywords	Language, cognitive ability, educational attainment, heritability, and gene-environment correlation
Date received	18 September 2022
Date approved	4 October 2022

Reference	**GDAC_2022_16**
Project title	Genomic Regulation of Pubertal Timing
Institution	Queen Mary, University of London
Application type	Data access / Genetics
Cohort	MCS
Abstract	Our health is not only determined by our genes, but also the environment we live in – which, in turn, can be impacted by ethnicity, economic status, and education, among other factors, also known as the socioeconomic determinants of health. Their impact has been well described in many health conditions affecting young people; however, this has not yet been studied in puberty. Puberty is the essential period of physical and psychological change from a child to an adult. Disorders of puberty are common, affecting 4% of children and are associated with poor long-term health outcomes, including obesity, type 2 diabetes, breast cancer, reduced fertility, and early menopause. This has a major impact on public health. This project aims to study the distribution of ethnicity and socioeconomic determinants of health in young people with disorders of puberty. This will help us understand if these factors play a role in disorders of puberty in the UK and improve our understanding of the causes of these disorders.
Keywords	Central precocious puberty, delayed puberty, pubarche, puberty, and socioeconomic deprivation
Date received	13 September 2022
Date approved	4 October 2022

Reference	**GDAC_2022_15**
Project title	Exploring the genetic architecture of developmental disorders using birth cohorts as controls
Institution	Wellcome Sanger Institute
Application type	Data access / Genetics
Cohort	NCDS and MCS
Abstract	About 1% of babies are born with a developmental disorder which affects mental or physical development, such as intellectual disability or

	congenital heart defects. Many of these disorders are at least partly genetic, often due to a rare change in a single gene. However, we also know that common genetic variants that influence cognitive ability, educational attainment and mental health traits in the general population affect risk of rare neurodevelopmental disorders too. We are exploring the genetic basis of rare neurodevelopmental disorders (NDDs) in two large cohorts of patients, Deciphering Developmental Disorders (DDD) and the Genomics England 100k Genomes project (GEL). We wish to use individuals from the 1958 birth cohort, the 1970 British Cohort Study, and the Millennium Cohort Study as population-based controls in these analyses, since most of them do not have NDDs. We will calculate genetic risk scores for different traits (including cognitive ability, educational attainment, and schizophrenia) and compare these between NDD cases and MCS participants. We will also make use of the trio data in MCS (children plus both parents) to investigate whether the genetic background of individuals' parents influences the risk of NDDs independently of the individuals' own genetic background ("genetic nurture"). Finally, we will explore how much of the risk of NDDs is explained by different types of genetic variants.
Keywords	Birth cohort controls and neurodevelopmental disorders
Date received	1 September 2022
Date approved	8 September 2022

Reference	**GDAC_2022_14**
Project title	Exploring the genetic architecture of cognitive, behavioural and mental health traits in MCS
Institution	Wellcome Sanger Institute
Application type	Data access / Genetics
Cohort	MCS
Abstract	About 1% of babies are born with a neurodevelopmental disorder, such as intellectual disability or autism. These disorders are at least partly influenced by genetic variation, and the genetic variants that influence them also influence cognitive ability, educational attainment, and mental health traits in the general population. Based on findings from cohorts of families with neurodevelopmental disorders, we suspect that some genetic variants may influence the trajectory of cognitive development rather than just the final outcome. Thus, this project will address these questions: • What are the genetic factors that influence cognitive ability, educational achievement, behaviour, and mental health traits in MCS children at different ages, and do these change across time? • To what extent are these genetic effects on children mediated via their parents' behaviour? • What is the relative role of genetics versus other factors such as socioeconomic status and parental education in influencing children's neurodevelopmental outcomes?
Keywords	Cognitive traits, behavioural traits, and mental health traits

Date received	1 September 2022
Date approved	8 September 2022

Reference	**GDAC_2022_13**
Project title	Genetic determinants of Olink proteomics in NCDS, and genetic correlations with proteomic data from other sources
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	The NCDS coordinating team recently acquired a large number of lab measures of proteins in blood samples collected from study participants at age 44 years. Studying determinants and consequences of differences in the amount of these proteins in the blood can offer many insights into the diagnosis, prevention and treatment of disease. A key step in using the information on the proteins is to ensure that the measures are valid, and have not been overly influenced by spurious factors, such as differences in the way samples were handled. One way to help validated protein measures is to compare the similarities between the genetic determinants of the proteins in the NCDS sample with the genetic determinants of the same proteins measured in samples elsewhere. This can be done by linking the genetic information on NCDS to the protein measures, and use these combined sources of information alongside publicly available summary information from other genetic studies of the proteins, e.g., findings published by the SCALLOP consortium (https://www.olink.com/our-community/scallop/).
Keywords	Proteomics, genetic correlations, GWAS, and Olink
Date received	16 August 2022
Date approved	8 September 2022

Reference	**GDAC_2022_12**
Project title	Genetic moderation of month of birth effects on educational attainment
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	MCS
Abstract	There is strong evidence that children's month of birth affects their performance in national (key stage) exams, with those older in the schoolyear (e.g., those born in September) performing better than those younger in the schoolyear (e.g., those born in August). We will check whether this discontinuity in school performance between August- and September-born exists also in the MCS cohort and examine whether this relationship differs for those with high and low genetic propensities to education, as well as how the effects change as children age. In other words, we will explore whether a high genetic predisposition for education exacerbates test score differences by month of birth.
Keywords	Gene-by-environment interplay, polygenic indices, social science genetics, and educational attainment

Date received	28 July 2022
Date approved	22 September 2022

Reference	**GDAC_2022_11**
Project title	The interplay of genetic and environmental influences on early child development
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	This project focuses on the topic of Gene-environment interactions (GxE), which refers to the question whether a genetic disposition moderates an environmental effect or whether the exposure to an environment changes the relevance of the genetic predispositions. In particular, the main question of the project is how genetic and environmental influences interact in influencing early child development. Concretely, we want to investigate (1) if and how genetic dispositions and the social origin shape the development of central child characteristics that are central for the further educational attainment, like cognitive skills or externalizing behaviour, over time and (2) whether certain genetic dispositions may modify the effect of environmental events.
Keywords	Externalising behaviours, gene-environment interaction, molecular data, polygenic score, and social inequalities
Date received	15 July 2022
Date approved	26 September 2024

Reference	**GDAC_2022_10**
Project title	The Genetics of Primary Lymphoedema and Lipoedema
Institution	St George's, University of London
Application type	Data access / Genetics
Cohort	NCDS
Abstract	We are seeking to identify mutations in genes which lead to the development of Primary Lymphoedema or Lipoedema. We have sequenced groups of patients with both diseases and will identify genes that are commonly mutated in these cohorts which are not as commonly mutated in a set of population-based controls. We plan to use the 1958 cohort sequencing data as our control set for this study.
Keywords	Genetics, primary lymphoedema, and lipoedema
Date received	13 July 2022
Date approved	8 September 2022

Reference	**GDAC_2022_09**
Project title	Understanding the role of adolescent cannabis use in the pathways to depression, anxiety, and suicidality
Institution	McGill University and Douglas Mental Health University Institute (Canada)

Application type	Data access / Genetics
Cohort	MCS
Abstract	The possible consequences of cannabis use on mental health remain poorly understood. While strong evidence supports associations between cannabis and psychosis, little is known about associations with depression, anxiety, and suicidality (suicidal ideation and attempt). Although cannabis use is highly influenced by genetic predisposition, previous studies have failed to account for pre-existing genetic vulnerability. This leaves unanswered the question of association vs. causality in the putative links with depression, anxiety, and suicidality, respectively. The general aim of the proposed project is to clarify the nature of the associations between cannabis use and depression, anxiety, and suicidality, respectively, taking advantage of genetically informed population-based designs. We aim to: 1) Investigate associations between cannabis use and depression, anxiety, and suicidality, adjusting for genetic confounders; and 2) Test whether individuals with genetic predispositions for a range of mental health problems are more at risk of depression, anxiety, and suicidality if exposed to cannabis (gene-environment interaction). Overall findings could improve mental health prevention/interventions on a possible modifiable risk factor, cannabis.
Keywords	Cannabis use and mental health
Date received	7 July 2022
Date approved	10 November 2022

Reference	**GDAC_2022_08**
Project title	The moderator role of genetic variants linked to nicotine metabolisation in gene CYP2A6 and their association with nicotine-devices initiation and use escalation between ages 14 and 17
Institution	Pennsylvania State University (United States)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Adolescents' electronic cigarettes (e-cigarettes) are a significant global health concern. E-cigarettes are perceived as safer and fashionable, and their various flavour options make them more appealing to adolescents who have never tried regular cigarettes before. E-cigarette use can promote nicotine dependence among youth, increasing the likelihood of frequent combustible cigarette use and decreasing quitting success. E-cigarette initiation and use escalation could be influenced, in part, by genetic variables. The CYP2A6 gene produces the CYP2A6 enzyme, responsible for nicotine metabolism. However, variations in the CYP2A6 gene might influence CYP2A6 nicotine metabolisation rates and influence smoking behaviours. For instance, individuals with genetic variations that make them faster metabolisers of nicotine might exhibit higher smoking

	frequency and deeper puffing behaviours than those who metabolise nicotine at slow or normal rates. Thus, faster nicotine metabolisers might be more frequent and intense smokers than slow or normal metabolisers even when controlling for social and environmental factors.
Keywords	Genetic variants, nicotine metabolisation, CYP2A6, nicotine-devices initiation, and nicotine use escalation
Date received	13 June 2022
Date approved	12 July 2022

Reference	**GDAC_2022_06**
Project title	Gaining Insights into the Causal Effects of Electronic Cigarette Use
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	MCS
Abstract	E-cigarettes can help people to stop smoking, but we do not currently know the long-term health effects of using e-cigarettes. We plan to use the available genetic data and data relating to smoking and e-cigarette initiation in the Millennium Cohort Study to investigate whether a person's genetic information can be used to predict vaping initiation (i.e., regularly using e-cigarettes or using e-cigarettes more than 100 times). We will conduct a genome-wide association study in which we will explore the link between genetic variants and e-cigarette initiation. This information can be used in a framework called Mendelian randomisation which helps us understand causal relationships. The results of this study will be used in further research which will help us to understand how using e-cigarettes may affect someone's health. It will also help us to understand whether e-cigarette initiation may lead to smoking initiation (known as the gateway hypothesis).
Keywords	E-cigarettes and long-term health effects
Date received	14 April 2022
Date approved	17 May 2022

Reference	**GDAC_2022_05**
Project title	Gene-environment interplay in early life cognitive and social-emotional development
Institution	University of York
Application type	Data access / Genetics
Cohort	MCS
Abstract	Genetic and environment factors, such as home environment, neighbourhood environment, and poverty, affect early child development. It is unclear whether the effects of genetics on early child development depend on environmental factors. This study aims to investigate whether the effects of genetics on cognitive and socio-emotional development depend on environment factors.

	We apply to use data on cognitive and social-emotional development, and a broad range of environment measures, from the assessment sweeps 1, 2, and 3 (ages 9 months, 3 years, and 5 years). In addition, we apply for polygenic scores (PGSs) to predict children's cognitive and social-emotional development. Using linear regression, we will investigate whether PGSs and environmental variables predict early child development. Using multiple regression, we will test whether environmental factors moderate associations between PGSs and early child development. We hypothesise that PGSs and environmental variables will predict early child development and we predict non-significant or small gene-environment interactions.
Keywords	Gene environment, early life cognitive development, and early life social-emotional development
Date received	1 February 2022
Date approved	26 April 2022

Reference	**GDAC_2022_04**
Project title	Childhood Adversity, Education, and Mental Health: Using Causal Mediation Methods to Estimate the Effects of Target Interventions Using Observational Data access / Genetics
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Exposure to adversities in childhood such as parental mental health problems, relationship difficulties or substance misuse are associated with an elevated risk of mental health problems. However, interventions aimed at breaking the link between these exposures and later mental health outcomes have only a limited evidence base. Moreover, interventions are very difficult and costly for researchers to evaluate, especially over longer time horizons. Our research will explore whether computer simulated interventions on environmental factors believed to mediate/transmit the effects of childhood adversities would result in improved mental health outcomes among affected children. Specifically, this research will investigate whether equalising multiple aspects of children's experiences at school between those with and without exposure to childhood adversities would lead to a reduction in later mental health inequalities. Although research into the possible health benefits of education has traditionally concentrated on attainment, our research takes a broader focus by also considering children's beliefs and feelings towards their school, teachers, and classmates as additional modifiable intervention targets. This work has the potential to flip current thinking on preventive mental health strategies by investigating whether interventions on the social environment rather than mental health symptoms directly might lead to meaningful improvements in population health.

Keywords	Causal mediation, childhood adversity, education; mental health, observational data, and target interventions
Date received	7 January 2022
Date approved	15 March 2022

Reference	**GDAC_2022_03**
Project title	National Child Development Study – Proteomics
Institution	University College London (UCL) – CLS
Application type	Samples request / Blood
Cohort	NCDS
Abstract	Proteomics analyses will be run on the blood samples collected from NCDS participants in 2002/3. This will substantially enhance NCDS and catalyse a step change in our understanding of the relationship between exposures from birth to midlife and their consequences for multiple physical and mental health disorders. It will provide high-dimensional biological information on these individuals during early midlife (age 42 to 44y), prior to the onset of most chronic disease, and at an age that is underrepresented in most cohorts, including UK Biobank (UKB). Embedding this technology within NCDS – with linkage to existing genetics and biomarker data, repeat measures of social and biomedical exposures, and pre-clinical and clinical disease outcomes will drive a major uptake in NCDS data use, including by large-scale international academic consortia aiming to understand the determinants of healthy ageing.
Keywords	Proteomics
Date received	4 January 2022
Date approved	10 February 2022

Reference	**GDAC_2022_02**
Project title	The role of sensory function for academic achievement in gene-environment correlations
Institution	University of Göttingen (Germany)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	What makes us successful when it comes to education? While there have been genetic and environmental explanation attempts, genetic and environmental factors are not independent from each other. For example, children with a genetic disposition towards higher educational attainment are more likely to enrol in advanced classes and therefore further increase their attainment. But what is it that makes these children select cognitively stimulating environments? We aim to investigate the role of childhood hearing and vision for educational attainment. First results suggest that a genetic disposition towards higher educational attainment is associated with better sensory function in childhood, resulting in better reading and school grades in

	adolescence. Analysing these associations longitudinally will help to determine time points for early interventions (such as correction of minor hearing and vision impairments). Since disadvantaged children show reduced hearing and vision, they could particularly benefit from interventions, resulting in a reduction of social inequalities.
Keywords	Academic achievement, gene-environment correlations, and sensory function
Date received	5 January 2022
Date approved	15 March 2022

Reference	**GDAC_2022_01**
Project title	Mental Health and Learning Difficulties: The Influence of Genes and the Early Home Environment
Institution	University of York
Application type	Data access / Genetics
Cohort	NCDS, BCS70, and MCS
Abstract	Learning comes easy to most children. But some have considerable difficulties in learning to speak, read, or do mathematics. These children are also at increased risk for lifelong mental health difficulties. But not all children with learning difficulties experience mental health difficulties. The reasons for such individual differences remain unclear but they are likely to be a result of genetic and environmental factors. Identifying pathways of vulnerability and resilience to various learning difficulties could help to reduce adverse mental health outcomes in some of the most vulnerable children. In the proposed project, we will investigate the extent to which genetic and environmental factors explain individual differences in children's mental health and learning difficulties. Specifically, 1. Which genetic variants are associated with children's learning difficulties? 2. What are the pathways through which vulnerability and resilience for mental health difficulties for children with learning difficulties are realised? (e.g., the early home environment).
Keywords	Genetics, intergenerational transmission, mental health, and special educational needs
Date received	4 January 2022
Date approved	13 January 2022

Reference	**GDAC_2021_19**
Project title	Testing the role of physical activity in promoting resilience against stress-related psychopathology
Institution	King's College London (KCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Mental health problems such as depression and anxiety affect one in eight children in the UK. Physical activity is associated with better

	mental wellbeing. However, it is not known whether engagement in regular physical activity can prevent the emergence of mental health problems in children, or whether this relationship is better explained by other genetic and/or environmental factors. This project will apply cutting-edge statistical methods to elucidate what effect physical activity can have on preventing the development of mental health difficulties in young people, and to disentangle its effects from other risk and protective factors, using data from large studies of children. By comparing results across different studies and statistical ways of analysing the information, it will be possible to draw more confident conclusions regarding whether interventions targeting physical activity could help to reduce the burden of mental health problems among young people.
Keywords	Physical activity and stress-related psychopathology
Date received	7 December 2021
Date approved	16 December 2021

Reference	**GDAC_2021_18**
Project title	DNA methylation signatures of psychedelic use
Institution	Karolinska Institute (Sweden)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Recent research suggests that use of so-called 'classic' psychedelics such as psilocybin ("magic mushrooms") and Lysergic acid diethylamide (LSD) may be associated with better physical health. For example, lifetime classic psychedelic use has been associated with lower odds of diabetes, heart disease and hypertension in the past year. Yet, the effects of psychedelic use on DNA methylation remain unknown. Here, using the National Child Development Study (NCDS), we carry out epigenome-wide association scans (EWAS) within each cohort dataset to identify psychedelic signatures on the blood methylome.
Keywords	DNA methylation signatures and psychedelic use
Date received	23 November 2021
Date approved	16 December 2021

Reference	**GDAC_2021_17**
Project title	Understanding the importance of genetic endowments in explaining the intergenerational elasticity of earnings
Institution	University of Cambridge
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Income across generations is persistent. Do genetic or environmental factors have a larger influence on this persistence? How do genetic and environmental factors interact? The nature vs. nurture debate is a topic that draws much attention in economics. However, only modest progress has been made so far due to data limitations. In this project,

	we plan to leverage three unique features in the NCDS to answer these questions: (1) detailed information on early-life investments, test scores, health, and educational attainment to understand the determinants of earnings; (2) lifetime earnings, measured using earnings data in all waves, as well as parental income at age 16, allowing measurement of the intergenerational elasticity of earnings (IGE); (3) the Genome Wide (GW) genotyping data can be used for Genome-wide association study (GWAS): subsets T1DGC, WTCCC2, WTCCC1, from which we can obtain a polygenic score to capture the relevant component of genetic endowment to predict earnings.
Keywords	Genes, human capital, intergenerational persistence, lifetime outcomes, and polygenic scores
Date received	19 November 2021
Date approved	15 March 2022

Reference	**GDAC_2021_16**
Project title	The Genetics of Occupational Status and Choice – A Genome-wide Association Study
Institution	Crest-ENSAE Paris (France)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Using data from more than 250 000 genotyped participants in another dataset (the UK Biobank), we investigate to what extent genetic differences are associated with occupational status and choice. In order to understand how generalisable our findings are, we want to test the predictivity of the identified genetic markers in the NCDS.
Keywords	Occupational status
Date received	14 November 2021
Date approved	16 December 2021

Reference	**GDAC_2021_15**
Project title	Association of polygenetic risk scores for clinical conditions, health behaviours and biomarkers with COVID-19: risk of infection and long COVID
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	NCDS and MCS
Abstract	Our genes can influence many of our characteristics, such as the predisposition to height and physical health. Polygenetic risk scores are used to summarise these effects with a higher score showing higher likelihood and a lower score lower likelihood of the characteristic. Many of these scores have already been produced in past research and can be applied to cohort studies with genetic information such as CLS cohort participants. This project aims to investigate the association of such Polygenetic risk scores for various diseases and biomarkers with the risk of COVID-19 infection and long COVID. The findings will tell us

	whether certain conditions and biomarkers are associated with a higher or lower risk of COVID-19 infection and long-term symptoms. This will help us understand biological pathways and inform whether there might be additional high-risk groups that had not been identified previously.
Keywords	PRS and long COVID
Date received	5 November 2021
Date approved	18 November 2021

Reference	**GDAC_2021_14**
Project title	Cross-cohort change in the genetic prediction of social outcomes
Institution	University College London (UCL) – CLS
Application type	Data access / Genetics
Cohort	NCDS and MCS
Abstract	The last 50 years have seen profound changes in socioeconomic outcomes—substantial increases in educational attainment, labour market changes, and widening income inequality. The DNA revolution has also occurred, and with it increasing evidence that genetic factors are important predictors of socioeconomic outcomes. This provides a powerful opportunity for this multidisciplinary project—to leverage genetic data across the globally unique British cohort studies to understand how the determinants of socioeconomic outcomes have changed. These determinants are cognitive performance, socioemotional skills, and mental and physical health measures.
Keywords	Genetic prediction and health
Date received	26 October 2021
Date approved	18 November 2021

Reference	**GDAC_2021_13**
Project title	Comparing the predictivity of life outcomes from genomic data to a ~250-word essay and teacher evaluations
Institution	Crest-ENSAE Paris (France)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	This project addresses the question of the predictability of important life outcomes (such as educational success, occupational status, income, etc.) from information available in childhood. Specifically, three types of information are compared: Genetic data, a short essay written at the age of 11, and an assessment of the respondent by a teacher, also at the age of 11. We want to find out how much these three different sources of data can tell us about a person's opportunities and prospects. For this, we use novel machine learning models to enable the best possible prediction.
Keywords	Human machine gene
Date received	23 October 2021
Date approved	7 December 2021

Reference	GDAC_2021_12
Project title	Analysis of Genomic Variation and Haemostasis Phenotypes
Institution	University of Washington (United States)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	The proposed research uses the B58C resources to investigate the association between genome-wide genetic variation and haemostasis traits, which includes plasma fibrinogen, fibrin D-dimer, von Willebrand factor, and tissue plasminogen activator. The purpose of the research is to better understand the genetic determinants of haemostasis and thrombosis. The research is being conducted by the CHARGE-TOPMed Hemostasis Working Group, which meta-analyses genome-wide association studies of haemostasis traits from dozens of studies. We will first impute the unified genotypes to a dense imputation panel and then will perform genome-wide association studies for each of the 4 traits. These summary data will then be meta-analysed with summary data from other studies and published in peer-reviewed journals.
Keywords	CHARGE and haemostasis
Date received	22 October 2021
Date approved	18 November 2021

Reference	GDAC_2021_11
Project title	Genetic susceptibility to sarcoma
Institution	Institute of Cancer Research (ICR)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Sarcoma is a rare cancer, representing approximately 1% of adult cancer diagnoses. In some cases, the risk of sarcoma development is increased by the presence of a disease-causing change in a sarcoma susceptibility gene. We plan to compare genetic sequencing data from individuals with sarcoma diagnoses to genetic sequencing data from control individuals without a cancer diagnosis. We aim to identify genes that may increase the risk of sarcoma development by conducting multiple individual case-control analyses using data from different sources. We will combine the findings of these analyses to generate our overall results.
Keywords	Genetic susceptibility and sarcoma
Date received	15 October 2021
Date approved	28 October 2021

Reference	GDAC_2021_10
Project title	Cardiovascular biomarkers in the 1970 British Cohort study
Institution	University College London (UCL)
Application type	Samples request / Blood
Cohort	BCS70

Abstract	Several cardiovascular biomarkers have recently raised interest in their use as prognostic factors. For instance, high-sensitivity cardiac troponin is a marker of cardiac damage used in NHS hospitals for the diagnosis of a myocardial infarction. However, that blood test has also remarkable value at population level, as mild elevations in healthy individuals are common and are associated with greater subsequent incidence of heart disease. The future of cardiovascular biomarkers research lies in the study of the causes and consequence of abnormal circulating levels using advanced epidemiological approaches within birth cohorts. My aim is to measure cardiac troponin and other important cardiovascular biomarkers in the stored blood samples from the 1970 British Cohort Study, in order to address several unanswered scientific questions, including but not limited to: • What are the reference values for cardiac troponin blood concentration in British middle-aged people? • What clinical and demographic variables modify the reference values for a normal troponin test and how? • What early-life factors are associated with elevated troponin at middle age? • What are the consequences of elevated troponin at middle age combined with other risk factors in terms of 5-year risk of myocardial infarction, stroke, and other cardiovascular disease events?
Keywords	Cardiovascular biomarkers
Date received	13 October 2021
Date approved	18 November 2021

Reference	**GDAC_2021_09**
Project title	Common genetic architecture of motor milestones and infant temperament
Institution	Birkbeck, University of London
Application type	Data access / Genetics
Cohort	MCS
Abstract	Motor milestones refer to learning to perform movements that involve the whole-body (like walking independently) or the hand and arm only (like stacking small blocks). Infant temperament reflects early development of personality and behaviour. Research has shown that motor development and infant temperament characteristics are sometimes associated with later mental health difficulties. While we know that mental-health conditions are partly inherited in families, we know little about the genetic influences on these indicators of behavioural development. Our project aims to understand the genetics of motor milestones and early temperament. We do this by testing the link between these infant characteristics and DNA differences that are commonly observed among people. We study this in infants who participated in multiple world-renowned studies including the Millennium Cohort Study (MCS). From this work we will derive information

	about genes and genetic summary scores for motor milestones and temperament that will inform future research.
Keywords	Motor milestones and infant temperament
Date received	7 September 2021
Date approved	20 September 2021

Reference	**GDAC_2021_08**
Project title	Association of polygenetic risk scores for clinical conditions, health behaviours and biomarkers with COVID-19: risk of infection and long COVID
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Our genes can influence many of our characteristics, such as the predisposition to height and physical health. Polygenetic risk scores are used to summarise these effects with a higher score showing higher likelihood and a lower score lower likelihood of the characteristic. Many of these scores have already been produced in past research and can be applied to cohort studies with genetic information such as CLS cohort participants. This project aims to investigate the association of such Polygenetic risk scores for various diseases and biomarkers with the risk of COVID-19 infection and long COVID. The findings will tell us whether certain conditions and biomarkers are associated with a higher or lower risk of COVID-19 infection and long-term symptoms. This will help us understand biological pathways and inform whether there might be additional high-risk groups that had not been identified previously.
Keywords	COVID-19, genetic variations, and long COVID
Date received	27 August 2021
Date approved	14 September 2021

Reference	**GDAC_2021_07**
Project title	CLS cohorts in the Polygenic Index Repository
Institution	National Bureau of Economic Research (NBER) (United States)
Application type	Data access / Genetics
Cohort	NCDS and MCS
Abstract	Our genes appear to partly influence many of the important 'outcomes' that social and health scientists seek to better understand, such as: our height and weight, our physical health, our mental health, our behaviour and even social outcomes such as how much education we obtain. One way to summarise information on genetics is to create a single score for each outcome—these are termed 'polygenic index' or 'polygenic risk scores.' Higher indexes mean someone has more genes which are linked with these outcomes. Researchers across many fields wish to incorporate genetics using these indexes, but there are many difficulties in doing so. This project seeks to add CLS studies to the Polygenic Index

	Repository. This means that we will create many of the indexes researchers will be interested in using, and the CLS will make them available to researchers for use in scientific research across many different disciplines.
Keywords	PRS
Date received	29 July 2021
Date approved	14 September 2021

Reference	**GDAC_2021_06**
Project title	The contribution of transmitted and non-transmitted risk alleles to child and adolescent anxiety and depression, and predictors of population-level youth mental health.
Institution	Cardiff University
Application type	Data access / Genetics
Cohort	MCS
Abstract	Anxiety and depression are common mental health difficulties that often start in late childhood or early adolescence. We know that the causes of anxiety and depression are complex and both genetic and environmental (for example, social) factors play a role. New research methods can use information from a young person and both of their biological parents to determine which genetic factors are passed on from parents to their child and what factors are not passed on from parents. These parental genetic factors that are not passed on can influence the family environment and are sometimes referred to as "genetic nurture". In this project, we aim to better understand the roles of genetic nurture, as well as genetic factors that are passed on to a child, in the development and course of anxiety and depression difficulties throughout childhood and adolescence. We also aim to further understanding of predictors of the developmental course of depression and anxiety, as well as possible explanatory factors for increasing trends in youth mental health.
Keywords	Anxiety, adolescence, depression, genomics, mental health, and time trends
Date received	26 July 2021
Date approved	27 September 2021

Reference	**GDAC_2021_05**
Project title	Trajectories of substance use and mental health problems: A genetically informed study
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	NCDS and MCS
Abstract	This study aims to investigate risk factors involved in the development of substance use problems, as well as the mental health consequences associated with early onset substance use problems. In particular, the aim is to assess genetic risks involved in the initiation and maintenance

	of use of different classes of substances (e.g., alcohol, cigarettes, cannabis), and how this genetic risk compares to the genetics associated with the manifestation of mental health problems. To identify genetic risks, a multi-trait genome-wide association study (GWAS) on a number of longitudinal phenotypes will be conducted, including those relating to substance use behaviours (age of onset of substance use, changes in substance use over time, frequency of substance use at different ages) and phenotypes relating to mental health outcomes at different developmental milestones.
Keywords	Substance use and genetic risks
Date received	14 July 2021
Date approved	28 October 2021

Reference	**GDAC_2021_04**
Project title	Exome sequencing of the Millennium Cohort study: rare variant effects on cognitive development and other phenotypes
Institution	Wellcome Sanger Institute
Application type	Samples and data access / Genetics
Cohort	MCS
Abstract	We will conduct DNA sequencing on all the offspring and some parents from the Millennium Cohort Study. This will generate a resource to enable research into the effects of rare genetic variants on a variety of traits in young people as they grow up. Specifically, we will sequence the genes of each individual, which make up ~1% of their total DNA sequence. We know that genetic variants within genes can have profound impacts on many traits, including medically relevant traits such as cholesterol levels or diabetes risk, as well as other traits such as height and cognitive ability. We will use the data to study the effects of these genetic variants on cognitive development, mental health, personality, and behavioural traits, and understand how these link to the genetics of severe neurodevelopmental disorders. The resource of DNA sequence data will be made available to other researchers to study many other traits.
Keywords	DNA sequencing and cognitive developments
Date received	9 June 2021
Date approved	29 June 2021

Reference	**GDAC_2021_03**
Project title	Differential Genetics Effect of Adiposity-Related Genetics through Childhood to Adulthood
Institution	University of Exeter
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Higher body mass index (BMI - a measure of weight corrected for height) may have different effects in early life compared to later life.

	However, studying the effects of BMI at different life stages is very difficult in randomised controlled trials. We would like to use the NCDS1958 study to investigate the role of childhood BMI independently of adult BMI using a novel genetic approach. This approach involves studying versions of genes that alter BMI at different stages of life. We would now like to look at these versions of genes in more detail in studies that have measured BMI in the same individuals across the life course so that we can be more confident about their effects. The National Child Development Study is perfect for these analyses because it has detailed long term measures of BMI and related developmental and disease traits. We will also explore how BMI and pregnancy history interact at different stages in the life course to modify risk of breast cancer. NCDS is one of few studies to provide the necessary details for this kind of analysis.
Keywords	Adiposity-related genetics, BMI effect on early life, BMI effect on later life, polygenic risk scores, and multi-variable mendelian randomisation
Date received	13 April 2021
Date approved	26 May 2021

Reference	**GDAC_2021_02**
Project title	Marital sorting by genotypes and phenotypes
Institution	Yale University (United States)
Application type	Data access / Genetics
Cohort	MCS
Abstract	We want to characterise how people form couples and how this affects child development. There is already some evidence on how individuals choose their partner on the basis of their education levels, and other observable features, indicating the presence of assortative mating. The evidence is almost non-existent in other dimensions and, in particular, on the assortative mating in terms of the genes that are related to the observable features on which assortative mating has been observed. Therefore, we will study assortative mating on the cognitive skills, non-cognitive skills, body mass index (BMI), height, and mental health levels (known as 'marital sorting by phenotypes') and also on the basis of the genes related to these phenotypes (known as 'marital sorting by genotypes'). We aim to provide evidence on how people's genes predict educational attainment, cognitive and non-cognitive skills, BMI, height, and depression, and in this way, how they relate to marital sorting. This is important, because as children inherit genes from parents, marital sorting can create innate advantages for some children and disadvantages for others. Therefore, marital sorting by genotypes and phenotypes could have important implications for inequality and intergenerational mobility.
Keywords	Marital sorting and child development
Date received	13 April 2021
Date approved	22 April 2021

Reference	GDAC_2021_01
Project title	The power of the environment: Environmental mediation of genetic liability
Institution	King's College London (KCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	The reasons why some children are bigger or smaller than others are complicated. We know, for example, that what and how much we eat can have impact on how heavy we are. Also, your parents body size has been found to be important. One important factor is how much we exercise and move our bodies. In this research project we want to find out how exercise can help you stay healthy, even though you might be at risk of having a bigger body. To do that we will use data from the Millennium cohort study, connecting measures of exercise, genetic information on body size, and measurements of body size during childhood. This will help us to understand how all these factors work together, and the degree to which exercise can help people to stay healthy.
Keywords	Exercise and body size
Date received	25 January 2021
Date approved	9 February 2021

Reference	MDAC-2020-0038
Project title	The G is in the E: how genes influence the effect of socioeconomic disadvantage on inhibitory control
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Being able to develop the capability of controlling attention, behaviour and thoughts, known as inhibitory control (IC) is important, not only because it is one of the main foundations that enables the development of more complex skills, such as intelligence and self-control, but also because low IC has been associated with criminal activity, substance abuse, and low levels of educational attainment. Being raised in a disadvantaged family and under a harsh parenting style have both been associated with child low IC. However, not every child raised in a disadvantaged family or under a harsh parenting style presents IC difficulties. This project aims at understanding individual differences in IC by studying if childhood's socioeconomic disadvantage has a causal effect on IC at 17 years old, and if this relationship is mediated by parental behaviour (e.g., parenting style). Additionally, it will explore if this effect is moderated by parents and child innate tendency to low inhibitory control. Understanding what causes the individual differences in IC can contribute to reducing inequalities, intergenerational transmission of poverty, and moreover, to improve well-being.

Keywords	Gene-environment interplay, inhibitory control, ADHD, socioeconomic status, parenting, and parent-child relationship
Date received	10 November 2020
Date approved	12 January 2021

Reference	**MDAC-2020-0037**
Project title	Neighbourhood effects in the gene-environment interaction on non-cognitive skills development and educational attainment across the life course
Institution	Crest-ENSAE Paris (France)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	This work focuses on individuals' development of non-cognitive skills (such as pro-social behaviour or self-regulation) since they are crucial determinants of success throughout life, including educational trajectories and attainment. Genetics account for the development of individuals' non-cognitive skills. Moreover, genetic predispositions can be suppressed or facilitated by specific environmental characteristics. Thus, we add to the literature on gene-environment interactions by taking into account the consequences of socio-economic deprivation within the neighbourhood on individuals' non-cognitive skills and educational trajectories. This project aims at using lifelong data from the 1958 birth cohort study on non-cognitive skills, education and neighbourhood deprivation, alongside genetic data, to investigate the following questions: 1. How do genetic predispositions for cognitive and non-cognitive skills affect non-cognitive skills development and educational outcomes throughout the life course? 2. How do neighbourhood deprivation and genetic predispositions interact in affecting non-cognitive skills development and educational attainment?
Keywords	Social science genetics, non-cognitive skills, educational attainment, and neighbourhood effects
Date received	9 November 2020
Date approved	29 January 2021

Reference	**MDAC-2020-0036**
Project title	Using the Millennium Cohort Study to perform family-based genome-wide association studies and environmental transmission of outcomes from parents to children
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	The causes of differences between people are complex and difficult to determine. These differences can be due to a mix of parenting, the environment, and genetics. Because people inherit both their genes and

	their environments from their parents, studies that include data from only offspring (that is, excluding data from parents as well) can be biased. In this project we will combine data from children, their mothers, and their fathers to investigate the genetic architecture of a range of health and social factors. By accounting for parental genetic data, we can ensure that our findings are not biased by environmental factors, something that has not previously been possible.
Keywords	Education, families, genetics, health, and longitudinal
Date received	2 November 2020
Date approved	12 January 2021

Reference	**MDAC-2020-0035**
Project title	Using a family-based study design to estimate the presence and impact of assortative mating in the Millennium Cohort Study
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Most people do not choose their partners randomly, but instead choose them based upon certain characteristics. For example, on average taller people tend to partner with other taller people. Where these characteristics have a genetic component, this process of choosing a partner will mean that people will not only be more similar in appearance to their partner than other people, but also more similar in terms of their genetics. Because people inherit both their genes and their environments from their parents, this selective partnering can bias studies that include data from only offspring (that is, excluding data from parents as well) and mean that they may provide inaccurate results. In this project we will combine data from children, their mothers, and their fathers to investigate how choosing partners (called assortative mating) can create inequalities between people which are maintained over generations.
Keywords	Education, health, genetics, families, and longitudinal
Date received	2 November 2020
Date approved	12 January 2021

Reference	**MDAC-2020-0034**
Project title	Using a family-based study design to estimate the causal effects of health and socioeconomic factors in the Millennium Cohort Study
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	MCS
Abstract	The causes of differences between people are complex and difficult to determine. These differences can be due to a mix of parenting, the environment, and genetics. Because people inherit both their genes and their environments from their parents, studies that include data from only offspring (that is, excluding data from parents as well) can be

	biased and provide inaccurate results. In this project we will combine data from children, their mothers, and their fathers. We will investigate how health factors such as BMI throughout life cause social outcomes such as education, and vice versa. We will use analytical methods that take advantage of the random allocation of genetic data at conception to estimate how health and education are causally related. Our project will provide new insight into the ways that health and educational differences between people may arise and be maintained over generations.
Keywords	Education, health, genetics, families, and longitudinal
Date received	2 November 2020
Date approved	12 January 2021

Reference	**MDAC-2020-0032**
Project title	Large-Scale Evaluation of the Effect of Rare Genetic Variants on Psychiatric Symptoms and Cognitive Ability – CNVs And Major Psychopathology (CAMP) study
Institution	Boston Children's Hospital (United States)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Copy number variants (CNVs) are rare genetic variants where a section of DNA is duplicated or deleted. CNVs are strongly associated with neuropsychiatric symptoms, such as depression, autism, psychosis, and intellectual deficits. However, current understanding of the impact of CNVs is limited because, due to their rarity, very large samples are needed to accurately test their effects. Moreover, previous studies have mostly focused on the most commonly recurring CNVs, leaving the majority unexamined. We aim to fill these knowledge gaps with a multisite, multidisciplinary, collaborative project that utilises existing datasets to achieve a very large sample size. We will model the effect of rare CNVs on neuropsychiatric symptoms, such as depression, autism, psychosis, and intellectual deficits. These models will be used to develop algorithms that can be utilised by researchers to formulate hypotheses, and by clinicians to estimate the contribution of CNVs in patients.
Keywords	Copy number variants, neuropsychiatric disorders, depression, autism, schizophrenia, intellectual performance, and cognitive ability
Date received	4 August 2020
Date approved	9 February 2021

Reference	**MDAC-2020-0031**
Project title	Family-based genome-wide association study of educational attainment and related traits
Institution	University of California (United States)
Application type	Data access / Genetics
Cohort	MCS

Abstract	Previous research has found that both genetics and the environment can affect an individual's level of education. Researchers have estimated the effects of specific genetic variants on education. However, these estimates capture not only the direct effect of the variants on the individual's education, but also the effects of the parents. To obtain the estimates we most care about, one needs to account for the parents' genetic data. We aim to obtain such estimates of the genetic variants' effects on education and related traits. We will use a method that properly accounts for the parents' genetic data. We will apply that method to MCS. MCS is one of the few datasets in the world with genetic data on a set of individuals and their parents. The estimates we obtain will enable a better interpretation of genetic studies of education and related traits. These estimates will also yield valuable insights into how genes interact with the environment.
Keywords	Education, polygenic score, genetic nurture, direct genetic effects, and indirect genetic effects
Date received	30 July 2020
Date approved	12 January 2021

Reference	**MDAC-2020-0030**
Project title	Cross-cohort change in the genetic prediction of health and socioeconomic outcomes
Institution	University College London (UCL) – CLS
Application type	Data access / Genetics
Cohort	NCDS, BCS70, and MCS
Abstract	This project will investigate links between genetic propensities and the following outcomes: body weight, smoking behaviour, socioeconomic position, and mental health. It will consider the following genetic propensities to the health outcomes (body weight, smoking behaviour and mental health), socioeconomic factors (e.g., education attainment), and cognitive function. It will use longitudinal data to examine how these links change across age. It will then compare estimates with analyses conducted separately in other birth cohort studies. This is because the importance of genetic factors is likely to depend on the society in which genes are expressed.
Keywords	Genetics, causation, longitudinal, health, social, and cognitive
Date received	28 July 2020
Date approved	17 December 2020

Reference	**MDAC-2020-0029**
Project title	What matters for who? The role of genes and cognitive and non-cognitive skills on intergenerational inequality
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	NCDS

Abstract	Many previous studies have shown a strong relationship between children's educational performance and their family background. However, these studies have been unable to tell whether this relationship is because of specific features of the child's educational path (e.g., the field of study chosen), child characteristics that are influenced by their parents (e.g., family social networks) or genetic differences between children. We will use cognitive (e.g., reading, remembering) and non-cognitive skill (e.g., behaviour, teamwork) data from the 1958 birth cohort study alongside genetic data to investigate two research questions: 1. How do people's skills develop with age, and to what extent do they explain associations between parental characteristics and offspring educational and employment outcomes? 2. Do genetic differences between children linked to differences in skills explain associations between parental characteristics and children's educational and employment outcomes?
Keywords	Education, inequality, genetics, socioeconomic position, and skills
Date received	7 July 2020
Date approved	3 September 2020

Reference	**MDAC-2020-0028**
Project title	A life-course perspective on the determinants of parenting
Institution	Duke University (United States)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	There is a wealth of research on parenting. Much of this research focuses on how parenting affects child development. There is less research on why parents parent their children the way they do. In particular, it is not clear how childhood experiences and characteristics of parents shape the way they parent their children once they become parents. Answering this question is important for two reasons. First, it can contribute to a better understanding of why parents in different families parent their children differently (for example, more or less strict). Second, it can contribute to a better understanding of why parents' behaviours sometimes pass on to their children. Doing this research requires studies that have followed individuals from when they were children up into parenthood. Such studies are rare. Here we propose to use a unique study – the 1958 cohort study – to investigate how experiences and characteristics of people influence their parenting once they grow up and have children.
Keywords	Parenting, life-course, children, longitudinal, and genetics
Date received	7 July 2020
Date approved	12 January 2021

Reference	**MDAC-2020-0027**
Project title	Behavioural problems of children: The role of parenting and genes
Institution	University of Amsterdam (Netherlands)

Application type	Data access / Genetics
Cohort	MCS
Abstract	Our study investigates why children behave in problematic and potentially dangerous ways. A particular focus is on the role parents play: Parents influence their children with the environment they create for them, as well as with the genes they pass on. Our study will yield better insights into how both of these factors matter and how they influence each other. The Millennium Cohort Study offers a unique opportunity to do so because it collected genetic data both from parents and their children, as well as detailed information about how children are raised and how they behave. The results of our research could identify ways that can prevent or reduce behavioural problems of children.
Keywords	Externalising behaviours, parenting, social science genomics, intergenerational transmission, trios, polygenic scores, and GWAS
Date received	31 August 2020
Date approved	12 January 2021

Reference	**MDAC-2020-0026**
Project title	Investigating the genetic causes of childhood short stature
Institution	Queen Mary, University of London
Application type	Data access / Genetics
Cohort	MCS
Abstract	The aim of our research is to better understand the underlying genetic causes of growth disturbances in children. Most of the differences in height between individual humans is determined by their genetic make-up. We aim to use the genetic data of the shortest 2% of individuals to look in detail at changes within genes that may potentially influence childhood growth. By comparing the genetic data of short individuals to that of children with normal growth, we hope to find patterns and traits that could cause growth delay. This is important because disorders of growth are very common and in severe cases, can be associated with long term health problems and reduced quality of life. This work will help us understand why some people are extremely short. We will also discover ways to treat children with medications that allow them to achieve a normal height, identify associated disorders and improve outcomes.
Keywords	Short stature, genetic causes, genotype-phenotype correlation, and next generation sequencing
Date received	15 July 2020
Date approved	12 January 2021

Reference	**MDAC-2020-0023**
Project title	Early parental investments, genetic endowments, and child development
Institution	Institute for Fiscal Studies (IFS)

Application type	Data access / Genetics
Cohort	MCS
Abstract	Early childhood is an important period for human development. During this period, parental nurturing interacts with the child's genes to shape his or her cognitive, social, and emotional development. Understanding the nature of these interactions is important to better understand the process of child development and inform policies that aim to enhance children's outcomes. In this project, we will model the process of cognitive, social and emotional development of children in the first five years of life, explicitly accounting for the role that child and parents' genes play in this process. The model will focus on three aspects of parental choice, namely the time spent by the child in formal and informal childcare, and the quality of home learning environment. We will study the extent to which the child's and parent's genes shape these choices and eventually affect children's outcomes. We will estimate the model using a large longitudinal survey of children born in the UK at the turn of the millennium.
Keywords	Cognitive development, socio-emotional skills, and genetic nurture
Date received	15 July 2020
Date approved	12 January 2021

Reference	**MDAC-2020-0022**
Project title	Understanding the intergenerational transmission of mental health
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	MCS
Abstract	Despite decades of research and intervention development, the causes of mental health problems are still poorly understood, and the burden of disease is rising globally. There is a well-established link between parent and child mental health, however, the mechanism through which this association acts is not fully understood. For example, maternal and paternal mental health problems (e.g., depressive symptoms) can affect their offspring both through genetic effects passed from parent to child and via the parenting environment. We propose to use a genetically-informed structural equation modelling approach to separate out direct genetic effects from effects of the family environment. These analyses require genetic information on mother-father-offspring trios, in addition to offspring mental health data. Disentangling these effects could inform future prevention and intervention strategies to improve child mental health and break intergenerational transmission of mental health problems.
Keywords	Intergenerational transmission, mental health, child mental health, and genetic epidemiology
Date received	15 July 2020
Date approved	16 October 2020

Reference	MDAC-2020-0021
Project title	Using genetics to understand growth in vulnerable babies and links with later respiratory health
Institution	University of Exeter
Application type	Data access / Genetics
Cohort	MCS
Abstract	Babies born with low birth weight (including those born prematurely and with rapid postnatal weight gain) often have poor respiratory health in the future, as we have shown using MCS data previously, but the underlying mechanisms are poorly understood. For example, it is not known whether babies who are growth-restricted in-utero are at the same risk as babies who are genetically small, or whether those who experience rapid post-natal catch-up growth in infancy are at the same risk as those who remain small. The rapidly growing amount of data on the genetics of birth weight and gestational duration, arising from efforts we are leading, offers an excellent opportunity to investigate mechanisms connecting early growth with later respiratory health. We will use genetic data from the MCS to compare babies who are genetically small with those who may be growth-restricted. We will examine whether growth-restricted babies are more likely to experience rapid catch-up growth, and whether they are at higher risk of poor respiratory health, than genetically small babies.
Keywords	Genetics, prematurity, intrauterine growth restriction, birth weight, catch-up growth, and respiratory health
Date received	15 July 2020
Date approved	16 October 2020

Reference	MDAC-2020-0020
Project title	Genes, parental behaviours, and child development
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Parents play a vital role in their children's development, through the time and money they spend with them, the way they interact with them, and the genes they pass them along. Until now, separating the contribution of each component has been difficult, due to the lack of data with genetic information on parents and children. In this project, using new data with genetic information on parents and children, we will study how parents influence their children's development, by disentangling what is the role of the genes they pass them on, the resources they provide, and also their parenting behaviours. The work in this project is supported by a prestigious ERC Consolidator Award on the role of genes and environments in shaping individual's life chances. It will contribute to a better understanding of the role of nature and nurture in shaping children's development, which is crucial to formulate policies to improve children's wellbeing and to reduce inequalities.

Keywords	Genes, polygenic score, child development, parental choices, genetic nurture, active gene-environment, correlations, evocative gene-environment correlations, and gene-environment interactions
Date received	7 September 2020
Date approved	17 December 2020

Reference	**MDAC-2020-0019**
Project title	The role of family background in children's educational attainment: A genetically sensitive, longitudinal cohort study
Institution	Duke University (United States)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Parents play a part in their children's school achievement, via at least three ways: first, via their financial resources. Second, via their parenting. Third, via their passing-on of genes to their children. This project will study how these factors combine to influence children's achievement. We will carry out two analyses. First, we will study how parents influence children's achievement through parenting and through the genes passed on to children. Specifically, we will test if parents' genes are linked with parenting behaviours that affect children's school achievement (e.g., reading to children). Second, we will take into account genetics when studying links between family background and children's academic and social skills, and later achievement. These analyses will form part of the "Early Years" chapter of the Institute for Fiscal Studies' "Deaton Review of Inequality in Britain." Our genetic analyses will focus on overall links across many people. They will not (and cannot) be used to predict outcomes for individual persons. The work in this project will contribute to a better understanding of how family background shapes children's lives.
Keywords	Childhood, parenting, attainment, genes, skills, and life-course
Date received	15 July 2020
Date approved	13 November 2020

Reference	**MDAC-2020-0018**
Project title	Childhood adversity and psychobiological trajectories: Examining social, cognitive, bio-behavioural, and immune pathways in youth
Institution	Lingnan University (Hong Kong)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Increasing evidence across countries indicates that childhood stress is associated with mental health development across adolescence. Exposure to childhood stress also has implications for biological systems and physical health. How exactly does early family environment shape these developmental and health outcomes in adolescence. An analysis of datasets involving response from both parents and children, on a wide range of family, school, community, and individual variables on

	children's cognitive, social, emotional, and physical development, will be conducted to examine how the childhood environment can get under the skin and disrupt multiple systems during adolescence. Results from this study will help us better understand the effects of childhood environment on adolescent development – psychologically, socially, cognitively, biologically, and the interaction between these domains and will inform early interventions for children in the future.
Keywords	Childhood adversity, youth development, mental health, biomarkers, and emotion regulation
Date received	15 July 2020
Date approved	28 October 2020

Reference	**MDAC-2020-0017**
Project title	Gene-environment correlations in mental health
Institution	Queen Mary, University of London
Application type	Data access / Genetics
Cohort	MCS
Abstract	It has been suggested that our genes don't just influence our mental health and wellbeing, but also have an effect on the types of environments we are exposed to (such as whether we smoke or the amount of social support we receive from friends and family). Until recently, the only evidence we had to support this view was from analyses of twins. However, we can now test this idea in unrelated individuals using their genetic data. We will explore whether genes that have been linked to mental health and wellbeing are also linked with the positive and negative experiences reported by the Millennium Cohort Study participants throughout their lives. Studying the link between genes and positive and negative experiences helps us to understand how genes influence our mental health and wellbeing and may uncover new targets for the treatment or prevention of mental health problems.
Keywords	Gene-environment correlation, psychology, mental health disorders, and rGE
Date received	7 September 2020
Date approved	15 October 2020

Reference	**MDAC-2020-0016**
Project title	Understanding the intergenerational transmission of risk for offspring mental health, cognitive and educational outcomes
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	MCS
Abstract	Parental risk factors are among the strongest early predictors of offspring mental health, cognitive and educational outcomes. This transmission of risk across generations hinders social mobility. Interventions targeting parents may thus appear promising. However, such interventions can only succeed if parental risks affect offspring

	outcomes via environmental pathways. Here, we propose to use genetically informed design to better characterise the intergenerational pathways underlying risk transmission. Rather than focusing on genetic influences per se, this project uses genetic information to evidence the impact of the environment, and in particular environments resulting from parental characteristics. The project will provide evidence on the relationship between parental risks and offspring mental health, cognitive and educational outcomes. On the longer term, it will help to design more efficient interventions by identifying developmentally relevant targets to prevent child adverse outcomes. This project is part of an international effort to better understand intergenerational risk for child outcomes and involves a range of stakeholders to ensure that findings are appropriately communicated and benefit children and families.
Keywords	Intergenerational transmission, mental health, cognitive ability, and educational outcomes
Date received	15 July 2020
Date approved	15 October 2020

Reference	**MDAC-2020-0009**
Project title	Using the 1958 birth cohort to test the population representativeness of controls in historical GWAS case-control studies
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Historical genome-wide association studies (GWAS) commonly used case-control designs whereby people with a disease (cases) were compared to disease free people (controls). Many of these people lay at the extreme ends of the disease spectrum; either being completely disease free or having serious disease. These GWAS designs assume that the selected control group are representative of the general population under study, such as the UK. If this is not the case, then GWAS results may be biased and provide estimates that are inaccurate or not generalisable to the broader population. In this project, we will investigate whether GWAS controls from the 1958 birth cohort are representative of other samples such as the UK Biobank.
Keywords	Education, health, genetics, socioeconomic position, and teacher reported characteristics
Date received	30 April 2020
Date approved	1 June 2020

Reference	**MDAC-2020-0006**
Project title	DNA methylation profiling in 1958BC samples to assess epigenetic responses to social and environmental cues in early life and over the life course across four UK population-based cohorts
Institution	King's College London (KCL)

Application type	Samples access / DNA
Cohort	NCDS
Abstract	Epigenetic mechanisms, including DNA methylation, are key regulators of gene function, and provide a potential interface for the interaction of the genome with the environment. There is emerging evidence that epigenetic mechanisms may mediate the effects of developmental and biological disruptions during the early years of life towards adverse age-related health outcomes. We secured ESRC/BBSRC funding to explore this hypothesis across four UK population-based cohorts. This funded project has generated genome-wide DNA methylation profiles in 240 whole blood DNA samples from the 1958 birth cohort (1958BC). We have recently secured additional funding to extend this project. In this current application extension, we seek approval to further profile DNA methylation in 1,368 additional whole blood DNA samples from the 1958BC. The newly generated epigenetic dataset, together with the original set of samples profiled by us, will result in altogether 1,608 whole blood DNA methylation profiles from the 1958BC. Possible benefits of the proposed research may include an improved understanding of molecular processes that underlie degenerative ageing processes and adverse health risks. The results have potential to inform lifestyle changes, improved strategies for health management, and gain insights towards development new preventative and treatment strategies.
Keywords	DNA methylation and Illumina Infinium HumanMethylationEPIC BeadChip
Date received	9 February 2020
Date approved	10 March 2020

Reference	**MDAC-2020-0005 – related to 2016_DNA_BCS70_1**
Project title	DNA methylation profiling in 1970BC samples to assess epigenetic responses to social and environmental cues in early life and over the life course across four UK population-based cohorts
Institution	King's College London (KCL)
Application type	Samples access / DNA
Cohort	BCS70
Abstract	Epigenetic mechanisms, including DNA methylation, are key regulators of gene function, and provide a potential interface for the interaction of the genome with the environment. There is emerging evidence that epigenetic mechanisms may mediate the effects of developmental and biological disruptions during the early years of life towards adverse age-related health outcomes. We secured ESRC/BBSRC funding to explore this hypothesis across four UK population-based cohorts. This funded project has generated genome-wide DNA methylation profiles in 236 whole blood DNA samples from the 1970 birth cohort (BCS70). We have recently secured additional funding to extend this project. In this current application extension, we seek approval to further profile DNA methylation in 264 additional whole blood DNA samples from the BCS70. The newly generated epigenetic dataset, together with the original set

	of samples profiled by us, will result in altogether 500 whole blood DNA methylation profiles from BCS70. Possible benefits of the proposed research may include an improved understanding of molecular processes that underlie degenerative ageing processes and adverse health risks. These signals have potential to inform lifestyle changes, improved strategies for health management, and gain insights towards development new preventative and treatment strategies.
Keywords	DNA methylation and Illumina Infinium HumanMethylationEPIC BeadChip
Date received	8 November 2019
Date approved	11 December 2019

Reference	**MDAC-2019-0934**
Project title	Predicting smoking histories from DNA methylation data to better understand the role of risky behaviours in health inequality
Institution	University of Essex
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Social and economic disadvantage affects people's health, leading to health inequality. Risky behaviours such as smoking and heavy drinking explain some but not all differences between the most and least advantaged people. However, research participants may not always be able to provide accurate smoking and alcohol data. In a previous project using UKHLS data, we show that differences in molecular indicators in the DNA, called methylation patterns, can be used to estimate smoking histories. If such estimates compare well with reported smoking in 1958BC, DNA methylation can provide an indicator, called a biomarker, of smoking across different studies. We will assess how this biomarker relates to questionnaire answers about smoking for different social groups and whether this adds to our understanding of health inequality. These methylation patterns could then help show how experiences like education shape DNA methylation which regulates our genes response to such circumstances and may partly drive different health outcomes.
Keywords	DNA methylation, smoking biomarker, and social health
Date received	21 February 2020
Date approved	24 March 2020

Reference	**MDAC-2019-0930**
Project title	Association of longevity with genetic variations in NPSR1
Institution	University of Jena (Germany)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	In a group of extremely old people (> 90 years) we have found genetic variations in the neuropeptide S receptor (NPSR1) gene that may influence the probability of reaching such a high age. However, the

	biological reasons for survival into advanced ages are currently unknown. We want to compare our results with data from the 1958 birth cohort and aim to identify health and disease factors that may confirm and explain the increased survival of people carrying certain NPSR1 variants. We will focus on factors linked with earlier death, such as heart disease, or functions of the immune system, and their relation to the NPSR1 gene.
Keywords	Neurotransmitter receptor, gene, and longevity
Date received	13 November 2019
Date approved	7 January 2020

Reference	**MDAC-2019-0927**
Project title	A genome and phenome-wide genetic association study of depression
Institution	University of Edinburgh
Application type	Data access / Genetics
Cohort	NCDS
Abstract	We wish to test the association between genetic risk of depression, which is fixed from birth, and its association with trajectories of child and adolescent development. Genetic risk of depression will be calculated using the genotypes already available in the 1958 birth cohort. The developmental traits of interest are those that are (a) measured on at least 3 occasions, and (b) of likely relevance, to physical, emotional and cognitive development and the emergence of common diseases of childhood. In addition, we seek to conduct a genome-wide association study (GWAS) of depression in the 1958 BCS, and chare the summary findings with the Psychiatric Genomics Consortium in a meta-analysis of similar studies – to see the accuracy of polygenic risk scores for worldwide and to contribute to the effort to better understand depression's genetic architecture.
Keywords	Depression and development
Date received	13 November 2019
Date approved	10 March 2020

Reference	**MDAC-2019-0916**
Project title	Gene-environment correlations in mental health
Institution	Queen Mary, University of London
Application type	Data access / Genetics
Cohort	NCDS
Abstract	It has been suggested that our genes don't just influence our mental health and wellbeing but may also influence the types of environments we are exposed to (such as whether we smoke or the amount of social support we receive from friends and family). Until recently, the only evidence we had to support this view was from analyses of twins. However, we can now test this idea in unrelated individuals using their genetic data. We will explore whether genes that have been linked to mental health and psychological wellbeing in previous studies, are also

	linked with the positive and negative experiences reported by the 1958 birth cohort study participants throughout their lives. Studying the link between genes and positive and negative experiences helps us to understand how genes influence our health and wellbeing and may uncover important targets for the treatment or prevention of mental health problems.
Keywords	Gene-environment interplay, gene-environment correlation, psychology, and mental health
Date received	3 September 2019
Date approved	9 January 2020

Reference	**MDAC-2019-0912**
Project title	Gene-environment interplay on reading skill and literacy promoting behaviours
Institution	University of Edinburgh
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Being able to read and write is related to life success. How well people can read and write is partly due to their genes. Because reading is a learned skill genes will interact with a person's environment. We will test genetic risk for dyslexia alongside measures of how a person is taught to read, and parental book reading, occupation. and income. This will help us understand whether there are aspects of the home and school environment that can lower a person's genetic risk of reading difficulties. We will check whether these are the same across childhood and youth, and if they affect adult outcomes like health. In the 1958 birth cohort we can test the association between all of a person's genes and their reading/writing abilities. We will combine these genetic results with those from the international Genetics of Language Consortium to improve knowledge of reading in different samples.
Keywords	Genetics, language, literacy environment, and reading skill
Date received	3 September 2019
Date approved	13 February 2020

Reference	**MDAC-2019-0019**
Project title	The Relationship Between Poor Health and Social Mobility
Institution	Imperial College London
Application type	Data access / Genetics
Cohort	NCDS
Abstract	The project will examine whether physical health affects social mobility – a change in the social and economic position of individuals – in the UK. Evidence suggests that health is related to social and economic position. But less is known about whether changes in social and economic position are affected by health. This project will use data from two studies: the 1958 National Child Development Study and Understanding Society to assess such relationship. Both studies have

	collected information on education, occupation, income, and health. In addition to interviews, these studies have conducted nurse-led assessments to gather health measurements and blood samples. These include measurements of blood pressure, height, weight, lung function, among others. The project will make use of biological measurements and genetic information to investigate the causal pathways that link different health conditions with the chances of moving social and economic position from one generation to another and of changing such position over a person's lifetime.
Keywords	Social mobility, inter-generational mobility, intra-generational mobility, and physical health
Date received	30 September 2019
Date approved	15 October 2019

Reference	**MDAC-2019-0007**
Project title	Investigating assortative mating and evocative gene-environment correlation in the 1958 cohort
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	NCDS
Abstract	There are large differences in education and health between people from privileged and disadvantaged backgrounds. However, the causes of these differences are complex and difficult to determine. One way that they may arise is due to the way that people select partners who are similar to themselves, and then create environments for their children based on their own characteristics. Because children inherit both their genes and their environments from their parents, genetic data offer unique ways to examine how inequalities arise. This project will use data from the 1958 cohort to answer two research questions: 1. Can selective partnering increase health and social inequalities over generations? 2. How are teachers' opinions of their pupils informed and how does this effect their education? This project will provide new insights into the ways that inequalities in education arise and are maintained over generations.
Keywords	Assortment, education, gene-environment correlation, genetics, and polygenic scores
Date received	3 May 2019
Date approved	10 September 2019

Reference	**MDAC-2019-0006**
Project title	Identification of novel breast cancer susceptibility genes in a Northern Ireland population
Institution	Queen's University Belfast
Application type	Data access / Genetics
Cohort	NCDS

Abstract	In the UK, 1 in 8 women will be diagnosed with breast cancer in their lifetime. Approximately 5–10% of these breast cancers are inherited (hereditary breast cancer), which may be caused by a change (a mutation) in a gene, which can be passed down through families. Mutations causing about one-third of hereditary breast cancers have been found in several genes. These mutations may mean that women who inherit them can be at a high risk of developing breast cancer in their lifetime. However, the genes contributing to the remaining two-thirds of hereditary breast cancers are unknown. This study aims to identify mutations in new genes that may increase the risk of breast cancer in women with a family history of the disease. Knowing which genes are involved in increasing the risk of developing breast cancer will be helpful for the future. We would then be able to carry out genetic testing to look for mutations in these genes. We may be able to offer these women more frequent screening (for example, with mammograms) to try and allow us to detect any cancers at an earlier stage. This might make it easier to treat them.
Keywords	Hereditary breast cancer, cancer genetics, and risk predisposition
Date received	20 March 2019
Date approved	25 April 2019

Reference	**MDAC-2019-0002**
Project title	Computation and Analysis of Polygenic Scores (PGS) in the National Child Development Study
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Recent research has found a large number of genes that relate to important social, economic and health outcomes. Information on several relevant genes can be combined into a single score, called a polygenic score. These simple scores can then easily be used in further projects. The first aim of this work is to calculate polygenic scores for various social, economic and health outcomes. Scores based on data from the National Child Development Study (NCDS) will be made available to researchers via METADAC or the UK Data Archive. This will be a valuable resource for investigating many policy issues where genetics plays a part. This new NCDS resource is designed to be used with comparable polygenic scores from two other studies: the Health and Retirement Study (HRS) and the English Longitudinal Study of Ageing (ELSA). The second aim is to demonstrate the usefulness of the newly computed scores with a showcase project. One of the polygenic scores is for genes that relate to 'years of education'. This project will compare genetic predisposition with actual outcomes and well-being in NCDS participants.
Keywords	Polygenic scores
Date received	1 April 2019
Date approved	17 October 2019

Reference	MDAC-2018-0025
Project title	Genetic Analysis of Individual Income
Institution	University of Amsterdam (Netherlands)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	In this project, we will study the relationship between genes and income. People with higher income tend to be more healthy and satisfied with life. Therefore, it is important to understand why some people have higher incomes than others. Previous studies suggested that individual differences in income are partly heritable, i.e., they are partly due to genetic effects. Our study will attempt to identify which specific genes are linked to individual differences in income and why these relationships occur. The results of our research will be useful to better understand how environmental factors (e.g., parenting, education) affect income, how genetic and environmental factors interact, and why higher income tends to be associated with better health.
Keywords	GWAS, income, and social science genomics
Date received	30 November 2018
Date approved	7 February 2019

Reference	MDAC-2018-0012
Project title	The Economic and Social Value of Health: Does childhood obesity hinder human capital development?
Institution	Imperial College London
Application type	Data access / Genetics
Cohort	NCDS
Abstract	The project will address an important set of research questions on the social and economic impacts of childhood obesity. It will use some of the most detailed longitudinal data sources available in the UK as well as innovative approaches to assessing causality and the links between health and social outcomes. In particular, the study will rely on two national cohort studies reflecting the lives of individuals born in 1958 and 1970, and on a local cohort study of children born in 1991-92 providing a unique set of information based on biomarkers, anthropometric measures, rich administrative data, and more traditional survey questions. The study will make use of biomarkers and genetic information to investigate the causal pathways that link children early life exposures and background socioeconomic status to their likelihood of developing obesity in young age and will examine the social and economic outcomes (e.g., educational attainment, cognitive skills, employment, earnings, civic participation, and social engagement) associated with childhood obesity.
Keywords	Children, obesity, human capital, and Mendelian randomisation
Date received	30 May 2018
Date approved	1 October 2018

Reference	MDAC-2018-0011
Project title	Investigating prevalence, type, and distribution of rare variants in ciliopathy genes in an unselected population
Institution	University of Zurich (Switzerland)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Ciliopathies are a group of recessive genetic disorders caused by dysfunction of primary cilia, little antenna-like structures present on most of our cells where they transduce signals from the environment to the cell. Clinical symptoms in this group of disorders can affect most organ systems individually or in combinations leading to syndromic disorders. Individual ciliopathy disorders are rare, but their precise prevalence is unknown since many affected individuals may not have received a diagnosis. The focus of this work is to provide a comprehensive evaluation of the prevalence, type, and distribution of rare sequence variants in ciliopathy-genes in an unselected population. This will help to better understand variants found in affected individuals and to estimate population carrier rates for this group of disorders.
Keywords	Ciliopathies, recessive disorder, variant interpretation, and carrier rates
Date received	14 May 2018
Date approved	29 May 2019

Reference	MDAC-2018-0009
Project title	Inheriting too many copies of the TPSAB1 gene causes allergies and food intolerance: How common is this genetic disease is in the UK?
Institution	Royal Manchester Children's Hospital
Application type	Samples access / DNA
Cohort	NCDS
Abstract	Together with colleagues at the National Institutes of Health, USA, we have recently discovered that patients with troublesome allergies, stomach trouble and loose and painful joints can sometimes have a high level of a chemical in their blood called mast cell tryptase (MCT) due to excessive number of copies of a gene called TPSAB1 in the patients' DNA. Using a group of American blood donors, we estimated that the frequency of this disease is between 4 – 6% of the population. The 1958 cohort offers a perfect population to determine how common this disorder is in an unselected British population. Four hundred DNA samples will be screened for the alpha-tryptase gene duplication. This will be performed using a new DNA sequencing method. It is called droplet PCR. We have just set it up and tested it out in our laboratory. This study will provide information as to how common this disease is and therefore highlight to doctors that it should be considered in patients complaining of the above symptoms.
Keywords	Allergy, mast cell tryptase, TPSAB1, and copy number variation
Date received	9 April 2018

Date approved	15 April 2019

Reference	**MDAC-2018-0005**
Project title	Biological embedding of childhood bullying, suicidal ideation, and depressive ideas
Institution	Douglas Mental Health University Institute (Canada)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Bullying is not uncommon, in fact approximately 40% of adolescents report being bullied by their peers. Prior research including evidence from the 1958 BBC indicates that being bullied in childhood is associated with poorer mental health including depressive symptoms and suicidal thoughts and behaviours in adolescence with persisting effect up to midlife. However, the underlying processes linking childhood bullying to suicidal ideation and depressive ideas into adulthood are unknown. Genes can modulate brain activity through molecular pathways and thereby influence mental health. This study will investigate 1) whether victims of bullying exhibited differential methylation patterns in their adult life and 2) whether specific methylation profiles are associated with suicidal ideation and depressive ideas. We aim to increase our knowledge of the biological embedding of bullying associated with suicidal ideation and depressive ideas.
Keywords	DNA methylation, peer victimization, epigenetic, and genome-wide
Date received	04 April 2018
Date approved	13 November 2018

Reference	**MDAC-2018-0005**
Project title	Peer bullying, epigenetics, and suicidal risk: prospective associations
Institution	Douglas Mental Health University Institute (Canada)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Suicide is one of the leading causes of death in youth worldwide. One of the key risk factors specific to the lives of children/adolescents that has attracted considerable public attention is bullying. Bullying is not uncommon: approximately 25% of adolescents' report being bullied by their peers. The goal of this project is to build new knowledge on the links between bullying and suicidality in the general population. First, our project will aim to determine whether bullying experiences increase the risk of suicide mortality. Second, since we know that genes can modulate brain activity through molecular pathways and thereby influence behaviour, we will investigate whether bullying has an effect on the regulation of genes and consequently on suicidality.
Keywords	DNA methylation, peer victimization, epigenetic, and genome-wide
Date received	29 March 2018
Date approved	21 May 2018

Reference	MDAC-2017-0027
Project title	Eczema remains common in adulthood: Results from 2 prospective British cohort studies
Institution	University of California (United States)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Eczema, an itchy autoimmune skin disease, is now one of the most common childhood diseases in industrialised countries. It waxes and wanes over time and has not been well studied beyond childhood. Using two longitudinal data sets representative of the UK general population (the 1958 and 1970 British cohort studies), our objective is to calculate how common active eczema is at multiple ages throughout childhood and adulthood and identify long-term patterns of disease activity. We will then test whether patient characteristics, including genetic factors, predict patterns of persistent, resolving, or late-onset disease. The anticipated benefits of this research include better information for patients on the likely progress of the disease. Data on how risk factors change over time could be used to design future preventative studies.
Keywords	Eczema, atopic dermatitis, epidemiology, and filaggrin
Date received	19 February 2018
Date approved	13 December 2017

Reference	MDAC-2017-0022
Project title	Associations between affective problems across the life course and cognitive function in midlife
Institution	University of Sussex
Application type	Data access / Genetics
Cohort	NCDS
Abstract	The primary aim of the proposed project is to investigate how affective problems, such as depression and anxiety, are linked with cognitive function in midlife, and to identify specific socio-behavioural and biological factors that may be involved in this relationship. Past research has shown that affective problems are associated with an increased risk of dementia and faster cognitive decline, but this research has been limited to old age samples with short follow-up periods. The current study therefore aims to investigate the association between affective problems over the entirety of the life course (from childhood to midlife) and cognitive function in midlife to further elucidate the temporal nature of this association over time, and to explore specific factors which may be involved within this relationship. We also plan to test the links between positive mental health and cognitive function and to test specific factors which may play a role in this relationship.
Keywords	Affective problems, mental health, cognition, genetics, and wellbeing
Date received	4 August 2017

Date approved	2 October 2017

Reference	**MDAC-2017-0021**
Project title	Investigating risk factors for continuities and discontinuities in emotional problems across the life course
Institution	King's College London (KCL)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Some childhood emotional difficulties persist into adulthood, while others do not. We plan to use the detailed information collected from members of the 1958 British birth cohort to identify people with different long-term patterns of emotional problems and explore individual and family factors that differ between them. Because emotional problems run in families, we will include markers of genetic risk, as well as factors such as relationships with parents and educational progress. We will also test whether the same factors can be helpful for children facing more severe stressors in their families, or who were bullied in childhood. If we can find factors that 'protect' some children with emotional problems from facing long-term emotional difficulties, this may give helpful pointers for treatment.
Keywords	Emotional problems, polygenic risk scores, and life course
Date received	16 August 2017
Date approved	3 October 2017

Reference	**MDAC-2017-0019**
Project title	Comparison of genome-wide DNA methylation between peripheral blood and lymphoblastoid cell lines
Institution	Temple University (United States)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Epigenetics is the study of the molecular switches by which genes are normally turned "on" or "off." There are three factors that influence whether these molecular switches are switched "on" or "off." The first is heredity (whether your parents had workable or defective switches, for example, and passed them on to you), the second is aging and the third is the environment (i.e., diet, exercise, pollution, etc.). We plan to compare the switching pattern of genes from individuals in the 1958BC with the switching pattern of multiple siblings from large families (using data from the 'CEPH families' project). This comparison will distinguish the effects of aging (because the 1958BC are all the same age) from the effects of heredity (because the siblings share half of their genes). Our hypothesis is that the remaining differences in switching pattern are the result of environmental factors. Ultimately, we hope this study will identify which environmental factors are most important in switching genes on or off.
Keywords	DNA methylation, lymphoblastoid cell lines, and peripheral blood

Date received	16 August 2017
Date approved	2 October 2017

Reference	**MDAC-2017-0018**
Project title	Investigating the genetic relationships between anxiety, depression, stressful life events, and cardiovascular risk factors and disease
Institution	King's College London (KCL)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	This study aims to use cardiovascular risk measurements and diagnosis, together with questionnaire data on mental and physical health in NCDS in two ways: 1. To discover and validate previous findings from large psychiatric genetics studies. These studies identified inherited genetic changes which may increase risk of depression. 2. To investigate shared genetic factors affecting mental and cardiovascular health (heart disease), as these conditions often occur together. The ultimate aim is to uncover biological pathways underlying the relationship between cardiovascular disease and depression. The proposed work will be achieved through analysis of genetic and health data, using existing methods. This research will assist in risk prediction, informing treatment, and forming a better understanding of the shared genetics between traits.
Keywords	Depression, polygenic risk scores, cardiovascular disease risk, and pleiotropy
Date received	14 August 2017
Date approved	19 September 2017

Reference	**MDAC-2017-0015**
Project title	Chronic Pain, Stress & Depression: Exploring Causal drivers
Institution	University of South Australia (Australia)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Depression and chronic widespread pain (CWP) frequently co-exist and interact in influencing health and well-being, often with devastating individual and societal consequence This project will investigate the relationship between chronic widespread pain and depression in mid-life, and whether chronic pain earlier in life affects depression risk later on. Furthermore, we will look to establish whether genetic variants which influence stress responses modify the relationship between chronic pain and depression. By investigating the role of genetic factors affecting stress responses as possible effect modifiers in this context, we hope to provide insights into more precise prevention strategies, and possibly, into biological pathways mediating the effects of chronic pain on depression.

Keywords	Chronic widespread pain, stress, depression, gene, and environment
Date received	14 August 2017
Date approved	10 July 2017

Reference	**MDAC-2017-0008**
Project title	Language and communication abilities as predictors of health and development during the life course – a genetic investigation
Institution	Max Planck Institute for Psycholinguistics (Netherlands)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Language is part of the human communication system and deeply embedded in social interaction and social behaviour. Beside mastering language milestones also social milestones play a crucial role during a child's development and has been linked to later behavioural, cognitive, social and health outcomes. However, the role of genes contributing to language, social communication and social behaviour that may mediate this relationship is not well understood, nor do we understand underlying causal relationships linking (early) childhood language and social behaviour to future development. This project aims to identify sets of genetic variants contributing not only to language but social communication and social behaviour and to use these sets as robust, confounding-free instruments to study genetic relationships with later life outcomes. Thus, this research can help predicting future health, wellbeing, education, and social life, and contribute towards the identification of developmentally sensitive windows in early/mid childhood.
Keywords	Comprehension, genetic association, health, lifestyle, reading, social behaviour, social cognition, social communication, social interaction, spelling, vocabulary, and writing
Date received	19 June 2017
Date approved	2 June 2017

Reference	**MDAC-2017-0007**
Project title	Gene-environment interactions in the female reproductive life course
Institution	University of Oxford
Application type	Data access / Genetics
Cohort	NCDS
Abstract	The female fertility window is bracketed by the first menstruation (menarche) at the start and menopause at the end. Although the timings of both menarche and menopause have been studied in terms of heritability and environmental influences, the two events have rarely been considered simultaneously in a genetic context. Furthermore, the reproductive history and choices of the studied population have not been taken into account as factors affecting the fertility timings. In this project, we propose to study the heritability of the length and position of the female fertility window. Additionally, we intend to combine the

	genetics with detailed reproductive history of the female participants to establish whether the timing of menopause is affected in part by their reproductive past (e.g., number and timings of pregnancies) as well as their reproductive choices (such as contraception and pregnancy terminations).
Keywords	Heritability, fertility, life history, menopause, and menarche
Date received	28 March 2017
Date approved	2 June 2017

Reference	**MDAC-2017-0003**
Project title	Genetic determinants of Human Reproductive Behaviour (HRB) – a collaborative meta-analysis
Institution	University of South Australia (Australia)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Recent genome-wide association studies have identified genetic variants linked to human reproductive behaviour such as: age at first birth and number of children ever born. The aim of this project is to extend this work, and to identify further genetic variants affecting indicators of reproductive success.
Keywords	Genome-wide analysis and reproductive behaviour
Date received	28 March 2017
Date approved	11 May 2017

Reference	**MDAC-2017-0002**
Project title	Personalised medicine in epilepsy
Institution	Centre hospitalier de l'Université de Montréal (CHUM) (Canada)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	We are building a pharmacogenomics study to decipher the pharmacoresistance component of epilepsy. We will compare 1000 pharmacoresistant epilepsy cases against 1000 good responder. All of this is being exome sequenced on Nimblegen. We need to build a control database in order to be able to perform a general comparison study and to be able to estimate the general burden of pharmacoresistance genes in the general population.
Keywords	Personalised medicine and epilepsy
Date received	13 February 2017
Date approved	2 June 2017

Reference	**2016_DNA_03**
Project title	Mapping of breakpoints in KLK3 deletions using exome sequence data, GWAS data, and laboratory validation
Institution	University of Bristol
Application type	Data access / Genetics

Cohort	NCDS
Abstract	Prostate specific antigen (PSA) can be measured in blood and is used in the screening, diagnosis, and monitoring of prostate cancer. The gene encoding PSA is called kallikrein-3 (KLK3). PSA levels vary: some people exhibit very low levels, and we found that some of these people may have deletions (parts of the gene are not present) or single-base mutations (changes of one 'letter' in the genetic code) in the KLK3 gene. These genetic changes may explain why these people have low PSA levels. It is possible that these individuals are more likely to have false-negative PSA test results. We think that ~1-2% of people may have KLK3 deletions (usually of one of two copies). We have determined the approximate location of these deletions, but not their exact location. We would like to use genetic sequence data and new genetic data from KLK3 to find the exact location of these deletions.
Keywords	Copy number variation, KLK3, deletions, prostate-specific antigen, exome sequencing, in silico CNV detection, and laboratory work
Date received	25 October 2016
Date approved	03 May 2016

Reference	**2016_DNA_02**
Project title	Epigenomics in the 1958 British birth cohort: an extension from pilot studies on social adversity
Institution	University College London (UCL)
Application type	Samples access / DNA
Cohort	NCDS
Abstract	It is widely accepted that early life influences shape our development, health, and behaviour across the life-course. Epigenetic mechanisms are increasingly implicated in these complex interactions. In a pilot study of 40 participants in the 1958 birth cohort we have contributed important evidence to this research by showing that DNA methylation (commonly used in epigenetic studies in human populations) has distinct associations with childhood socio-economic position and abuse. Building on this work we now plan to collaborate with several other cohorts to address questions that advance understanding on what aspects of our environment impact upon gene regulation, how our environment and way of living become embodied in human biology, over what timeframe and with what degree of persistence and how social and biological inequality may influence development and health.
Keywords	DNA methylation, prenatal tobacco exposure, childhood adversity, adult socio-economic position, child to adult health and growth outcomes, and behaviour
Date received	05 August 2016
Date approved	01 August 2016

Reference	**2016_DNA_01**

Project title	Genome-wide association study of amyotrophic lateral sclerosis
Institution	National Institute on Aging, NIH
Application type	Samples access / DNA
Cohort	NCDS
Abstract	We are undertaking laboratory experiments to find genes that may cause amyotrophic lateral sclerosis (ALS). ALS is a fatal condition that kills approximately 1,500 UK residents every year. Individuals with this condition become weak due to muscle wasting and typically die with two to three years of the onset of their symptoms. There is currently no cure for this disease. Knowing the genes that cause this disease may help us to design therapies to help patients. A person who does not have the disease being studied is known as a control subject. We will use the DNA obtained from the British 1958 birth cohort as control subjects in our genetic study. Results from control subjects will be compared to results from British ALS patients. The availability of this control information is crucial to the success of our project and increases our power to find the causative genes.
Keywords	Genome-wide association study, amyotrophic lateral sclerosis, and genetic risk
Date received	05 August 2016
Date approved	18 March 2016

Reference	**2016_DATA_16**
Project title	Cognitive capability across the lifespan: the interplay between the APOE gene and environmental factors
Institution	University of Sussex
Application type	Data access / Genetics
Cohort	NCDS
Abstract	As part of a PhD project at the University of Sussex into cognitive function over the lifespan, funded by the Alzheimer's Society, we wish to investigate whether one of the key genetic risks for late life dementia, the Apolipoprotein E gene (APOE), can be modified by environmental factors. It has been shown that carriers of a certain variant of this gene (the ε4-allele), are at increased risk of developing Alzheimer's disease and demonstrate poorer cognitive ageing. It has been reported that higher engagement in cognitively stimulating activities can improve an individual's likelihood of good cognitive ageing. We wish to explore how early life activities play into this interaction. This project examines how the APOE gene interacts with early and late life cognitive and leisure activities (e.g., reading books, physical activity, etc.) to affect cognitive performance over the lifespan.
Keywords	Ageing, Alzheimer's disease, cognition, cognitive activity, education, socioeconomic status, and Apolipoprotein E.
Date received	25 July 2016
Date approved	15 March 2017

Reference	2016_DATA_15
Project title	Beta-defensin copy number variation and obesity
Institution	University of Leicester
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Obesity is a common trait. People's weight is dependent both on their genetic makeup and lifestyle choices, such as amount of exercise. Understanding the variation in genes that is responsible for the genetic component has identified many variants in the DNA, each of which contribute a small amount to an individual's weight (usually measured as body mass index, BMI). In this study, we have new functional evidence that the beta-defensin genes are important in modulating an individual's weight. However, these beta-defensin genes are usually ignored by current genetic studies because they are in a region of the human genome that is not well-measured by most approaches. Beta-defensins show variation in the numbers of genes in the genome. We wish to investigate whether there is a correlation between the number of beta-defensin genes an individual has and their BMI.
Keywords	Obesity, BMI, and copy number variation
Date received	21 July 2016
Date approved	14 December 2016

Reference	2016_DATA_14
Project title	Causal role of fatty acids metabolites in reading and spelling measures
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	NCDS
Abstract	We have observed that higher levels of mono- and poly-unsaturated fatty acids are correlated with higher scores for reading and spelling tests in 7-year-old children. Previous evidence supported that some of these fatty acids participate in brain development, maintenance, and function. In this project we will investigate whether modifying these fatty acids can result to a change in education performance measures. Our study will make use of an alternative to randomised controlled trials, based on the natural randomisation of genetic variation in the population. For this, genetic variants previously linked to fatty acids are tested for association with academic performance measures. Since the genetic variants are not associated with the other measures that are possible to influence the link between academic performance measures and fatty acids, such as socioeconomic variation, any association identified can be considered as causal.
Keywords	Fatty acids, education, and Mendelian randomisation
Date received	23 June 2016
Date approved	23 October 2017

Reference	2016_DATA_12

Project title	Advanced modelling of genomic data
Institution	University of Southampton
Application type	Data access / Genetics
Cohort	NCDS
Abstract	With the advent of new and cheap technologies for reading the human genome, we are now able to collect data from thousands of patients. However, making sense of this incredible amount of data is not always easy and old analytical tools are no longer effective. Our project involves the application of modern mathematical computational algorithms, called machine learning algorithms, to genomic data. These mathematical tools can learn patterns by reading genomic data and identify series of mutations that might be crucial in defining a particular condition. However, in order to extract this kind of information, we also need data from individuals that are not presenting the same investigated clinical condition. The 1958 birth cohort study is an excellent unaffected (control) dataset that will make possible the application of machine learning algorithms to many different complex and rare disorders, such as cancer, diabetes, and chronic diseases.
Keywords	Next generation sequencing, machine learning, and genomics
Date received	17 May 2016
Date approved	23 September 2016

Reference	**2016_DATA_10**
Project title	Prediction of complex traits using whole-genome methods
Institution	Michigan State University (MSU) (United States)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	We will use this data set in combination with other data sets to assess the effects of sample size, marker density, and of the estimation method on the prediction accuracy of Whole Genome Regression methods when applied to a variety of complex human traits, including anthropometric and health-related traits.
Keywords	Complex traits, Whole Genome methods, and prediction
Date received	06 May 2016
Date approved	19 October 2016

Reference	**2016_DATA_09**
Project title	Investigating oligogenicity and genetic modifiers in ciliopathies
Institution	University of Zurich (Switzerland)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Joubert syndrome (JBTS) is a genetic disorder defined by a specific brain malformation accompanied by intellectual disability and problems with movement coordination. In addition, the clinical presentation can be very variable since some affected patients additionally develop

	disease of the retina, kidneys, and liver. JBTS can be caused by mutations in both copies of one among 30 different genes, but because the genetic cause has still not been identified in one third of affected families, alternative inheritance mechanisms have been proposed, whereby single mutations in two or more different genes could also cause the disorder ("digenic/oligogenic inheritance mechanism"). Moreover, the variability in clinical presentation is thought to arise from combined effects of variants in multiple genes ("genetic modifiers"). The focus of this work is to systematically investigate these hypotheses by comparing the proportion of JBTS patients with mutations in multiple genes, to that in healthy control individuals.
Keywords	Ciliopathies, Joubert syndrome, phenotypic variability, genetic heterogeneity, recessive disorder, oligogenicity, and genetic modifiers
Date received	06 May 2016
Date approved	26 August 2016

Reference	**2016_DATA_08**
Project title	The Role of Protein-altering Genetic Variation in the Pathogenesis of Silver-Russell Syndrome
Institution	King's College London (KCL)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Silver-Russell syndrome (SRS) is a developmental disorder with varied features. These include very small size at birth, poor subsequent growth, a triangular face, asymmetrical body, and a range of other possible minor problems. The precise cause of SRS is unknown. However, previous investigations suggest some cases of SRS are caused by genetic mutations, meaning errors which provides the blueprint for normal development. In this project we aim to test whether two types of genetic errors could cause SRS: (a) a single mutation with a large effect, or (b) many smaller-effect mutations in regions of DNA which are important for development. To do this will study the DNA sequences of 21 SRS patients. We will compare these sequences against 1,000 unaffected people to help pick out the mutations which may cause SRS from the many non-harmful genetic errors which will also be present.
Keywords	Silver-Russell syndrome, developmental disease, Mendelian disease, and whole exome sequencing
Date received	18 April 2016
Date approved	17 June 2016

Reference	**2016_DATA_06**
Project title	Identification of novel risk factors and causative mutations of neurodegenerative syndromes
Institution	University College London (UCL)
Application type	Data access / Genetics
Cohort	NCDS

Abstract	Clear evidence for a genetic component to many neurodegenerative diseases comes from the fact that relatives of patients have an increased chance of getting the disease themselves. However, few cases have an identified "spelling mistake" mutation in the known disease genes. We believe there are further disease-causing genes to be identified. Next-generation sequencing technology now allows us to analyse all the protein coding regions of an individual's genome in a single experiment (called exome sequencing). By using exome sequencing, we aim to identify novel genes/mutations that cause disease and genetic changes that modulate risk. Using the control data from the 1958BC will allow us to more accurately determine if newly identified mutations/genes are involved in the disease process. New mutations identified would help lead to a better understanding of the genetic makeup and disease mechanisms of diseases, which could lead to the development of more appropriate strategies to combat them.
Keywords	Neurodegeneration, genetic risk factor, mutation, Alzheimer's disease, Parkinson's disease, fronto-temporal dementia, and dementia with Lewy Bodies
Date received	24 March 2016
Date approved	17 June 2016

Reference	**2016_DATA_05**
Project title	Investigating Hereditary Cancer Predisposition – a combined genomics approach
Institution	University of Cambridge
Application type	Data access / Genetics
Cohort	NCDS
Abstract	The future of diagnosing cancer earlier, especially in patients with strong family history for cancer, is in the identification of further inherited genes relating to cancer development. The genes identified as predisposing (or increasing the likelihood) of cancer, could then be screened within the family to alert relatives of their increased risk, whilst also potentially assisting with therapy of the original patient by providing more tailored treatment options. We will use the ICR1000 data from the 1958 birth cohort, collected with similar technologies and population background to our patient samples, as a standard reference dataset. We will compare it with our data and identify the most likely predisposing genes in our sample sets.
Keywords	IHCAP
Date received	4 March 2016
Date approved	17 June 2016

Reference	**2016_DATA_04**
Project title	Identification of Genes Contributing to Neurodegenerative Diseases
Institution	University of Massachusetts (United States)
Application type	Data access / Genetics

Cohort	NCDS
Abstract	The requested exome data will be used as controls in a study to identify genes contributing to ALS, Parkinson's disease, and other neurodegenerative diseases.
Keywords	Case-Control, exome sequencing, and neurodegenerative diseases
Date received	18 February 2016
Date approved	26 August 2016

Reference	**2016_DATA_03**
Project title	Renal Cancer Genetic Association Study Following Different Strategies
Institution	University of Cambridge
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Understanding the role that genetic variants have on the onset and course of a complex disease like renal cancer is key to achieve a better treatment for the patients suffering from this disorder. Renal cancer (as most cancers) is a highly genetically heterogeneous disorder, and in this sense, we intend to tackle with a different approach, use a cohort of renal cancer patient with extreme phenotype, which should reduce considerably the heterogeneity of the genetic component involved in this subgroup of the disease. This approach has already been used successfully in the study of the genetic component of other human traits and will surely yield new insights into the genetic component of renal cancer.
Keywords	Renal cancer, genetics, association study, and complex disease
Date received	17 February 2016
Date approved	11 May 2016

Reference	**2016_DATA_02**
Project title	A genome-wide interaction study on life course health effects from childhood immunomodulation to micropathogens
Institution	Johns Hopkins University (United States)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Childhood immunisation is a powerful tool in reducing deaths related to infectious diseases. Several lines of evidence now suggest vaccination may confer additional unintended health benefits. While the biological mechanisms and range of these vaccine health effects have not been fully clear, chronic infection is known to trigger non-communicable diseases through inflammation and other immunological processes. Using phenome-wide data from the CLS cohorts, we recently characterised specific health changes throughout a person's life, following certain types of vaccination and vaccine-preventable illness. The goal of this study is to now pinpoint the genetic drivers of these non-specific health effects and identify how external factors such as lifestyle and environment may influence these pharmacogenetic effects.

	The research will use a combination of high-dimensional, causal-based inference methods, and computational genomics to analyse this data.
Keywords	Genome-Wide Interactions, vaccine, vaccine-preventable illness, non-communicable disease, and life-course
Date received	11 February 2016
Date approved	18 March 2016

Reference	**2016_DATA_01**
Project title	Germline susceptibility to cancer
Institution	Memorial Sloan Kettering Cancer Center (United States)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	We propose to use this data as a set of convenience controls to compare against sets of cases with cancers. We will use this for common cancer types such as breast, gynaecologic, urologic, and gastrointestinal. The basic premise is that the genotype data generated from the ICR1000 UK exome series will act as a control set when matched properly for genes covered and when adequate quality matching has been performed.
Keywords	Cancer genetics, genomics, and bioinformatics
Date received	10 February 2016
Date approved	18 March 2016

Reference	**2016_DNA_BCS70_1**
Project title	DNA methylation profiling in 1970BC samples to assess epigenetic responses to social and environmental cues in early life and over the life course across four UK population-based cohorts
Institution	King's College London (KCL)
Application type	Samples access / DNA
Cohort	BCS70
Abstract	The epigenome provides a mechanism of interaction between the genome with the environment, and we hypothesise that early life stimuli and exposures over the life course leave an epigenetic mark. The proposal aims to explore epigenetic profiles in up to 200 whole blood samples from the 1958 birth cohort, to contribute towards a large discovery analysis of 4,000 samples from four British cohort studies in order to identify epigenetic signatures of early life experience and exposure to social, environmental, and biological stimuli over the life course, linking findings to changes in physical and cognitive function during ageing. Each cohort captures a range of early life experiences, longitudinal health measures and lifestyle questionnaire data from adult life and has available DNA samples. The proposal will address four research aims, with overall purpose of identifying epigenetic signatures of early-life events and exposures over life-course that may have impacts on functional health trajectories.
Keywords	DNA methylation and Illumina Infinium HumanMethylationEPIC BeadChip

Date received	8 November 2019
Date approved	11 December 2019

Reference	**2015_DATA_01**
Project title	Rare Variants in Severe Obesity
Institution	University of Cambridge
Application type	Data access / Genetics
Cohort	NCDS
Abstract	We are interested in findings genes involved in obesity. We do this work by studying children who develop severe obesity from a young age. We have used a sophisticated technology that allows us to look at thousands of genes at the same time. However, the challenge is working out how to make sense of this huge amount of information. One of the most useful ways to start is to compare the genetic information we have with similar information obtained in people who are healthy. This will allow us to find genes that are specifically involved in severe obesity, information which could help us to find new treatments for obesity and diabetes in the future.
Keywords	Obesity and genetics
Date received	11 December 2014
Date approved	2015

Reference	**2015_DATA_02**
Project title	Identification of SNP variants and CNVs associated with different congenital heart disease phenotypes
Institution	McGill University (Canada)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Congenital heart disease covers a wide spectrum from small defects, which may be totally asymptomatic and compatible with a normal lifespan, to more severe forms which require urgent intervention. Many defects are possible, but most defects either obstruct flow of blood in the heart or in vessels near to it or cause blood to take an abnormal route through the heart. More rarely only one ventricle may be present, or the right or left side of the heart has failed to form properly (hypoplastic heart). Significant amounts of blood shunting from right to left without traversing the lungs causes cyanotic congenital heart disease. With access to genetic data, we try to identify genetic causes for this disease.
Keywords	Variants, SNP, disease phenotype, and congenital heart disease
Date received	8 January 2015
Date approved	2015

Reference	**2015_DATA_03**
Project title	The epidemiological relationship between low mood and autoimmune phenotypes across the lifespan – a genetic perspective

Institution	King's College London (KCL)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	We propose using genetic data to calculate 'risk profiles' in NCDS participants. We can then use these to investigate the contribution of genetics to the disorders individuals develop, across their lifetimes.
Keywords	Genetic profile risk score, autoimmune, mood, malaise, depression, comorbidity, and longitudinal
Date received	13 January 2015
Date approved	2015

Reference	**2015_DATA_04**
Project title	Genetic susceptibility for health and illness: Collaborative Genetic associations and gene-environment interaction analyses
Institution	University of South Australia (Australia)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Lifestyle factors affect the effects of genetic factors on the patterns of growth and development, and how those translate into a future risk of diseases. There is great interest in establishing related gene-environment interactions and how these operate across the life-course. In this large-scale collaborative genetic epidemiological project, we propose to use information collected from the 1958 British birth cohort to investigate genetic influences, gene-environment interactions, and life-course influences affecting patterns of growth, health, and diseases risk. This work is important not only in helping us to understand how genes affect health and disease, but also in giving a better understanding on how by making changes to our lifestyles we can attenuate negative effects caused by possible genetic susceptibilities.
Keywords	Growth, health, cardiovascular disease, diabetes, metabolic risk markers, gene-environment interaction, and genetic association
Date received	23 February 2015
Date approved	30 September 2015

Reference	**2015_DATA_05**
Project title	Genetic associations and gene-environment interactions affecting cognition, behaviour, and mental health: Collaborative analyses
Institution	University of South Australia (Australia)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Genetic factors affect our behaviour, cognitive ability and mental health, and there is evidence to suggest that lifestyle factors can modify related influences. In this large-scale collaborative genetic epidemiological project, we propose to use information collected from the 1958 British birth cohort to investigate genetic influences and gene-

	environment interactions affecting behaviour, cognition, and mental health. This work is important not only in helping us to understand how genes affect our behaviour and health, but also in giving a better understanding on how we can attenuate possible negative effects caused by genetic susceptibilities by making changes to our lifestyles.
Keywords	Cognition, behaviour, mental health, gene-environment interaction, and genetic association
Date received	21 May 2015
Date approved	30 September 2015

Reference	**2015_DATA_06**
Project title	Is skin tone associated with vitamin D and associated outcomes? A Mendelian randomisation study
Institution	University of Bristol
Application type	Data access / Genetics
Cohort	NCDS
Abstract	This project will investigate whether individuals with fairer skin have higher vitamin D levels. We will then investigate whether this variation in skin tone and any variation in vitamin D has an effect on BMI and blood pressure. We will predict variations in individuals skin tone using small variations in individuals' genetic makeup that have been shown previously to be associated with skin tone.
Keywords	Skin tone, ethnicity, and vitamin D
Date received	21 May 2015
Date approved	28 July 2015

Reference	**2015_DATA_07**
Project title	Genetic Studies of Methylation Genes and Cognitive Phenotypes
Institution	University of Nottingham
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Mutations within genes encoding methylation related proteins are known to cause or predispose individuals to developing brain diseases including familial forms of dementia and neurodevelopmental phenotypes. Common genetic variants within such genes have also been associated with differences in cognitive ability. We are currently using the Oxford Project to Investigate Memory and Ageing cohort to explore genetic variation with cognitive performance and biochemical measures. We want to conduct replication/follow-up studies using the NCDS datasets. Access to NCDS data is being sought for the: special License Biomedical Data set; sweep 8 cognitive assessments and health related variables, and; the GWAS resource. We wish to link these datasets together with the aim to explore genotype, cognitive performance, and other factors which have the potential to affect health status (for

	example, biochemical measures, physical activity; general health, and diet), medication use, and mental health measures.
Keywords	Methylation, genes, and cognition
Date received	9 June 2015
Date approved	29 July 2015

Reference	**2015_DATA_08**
Project title	NUT Midline Carcinoma Research Program
Institution	Telethon Kids Institute (Australia)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	NUT Midline Carcinoma (NMC) is the name of a very rare but lethal cancer that affects both children and adults. Even though NMC tumours initially respond to chemotherapy, all patients quickly relapse and there is no effective treatment. There are thus no survivors of this dreadful disease and patients usually succumb within a few months from diagnosis. We are analysing the biology of the disease to try and find clues to better treatments. To do this we have used state of the art DNA sequencing technologies, however we are limited in our ability to analyse the data because of the scarcity of samples from NMC patients. In order to validate our results, we need to compare them to samples from normal individuals (i.e., people that do not have cancer or other known diseases). The purpose of this application is to request these normal samples from the EGA study EGAS0000971 which is a UK study that has studied a large number of normal individuals.
Keywords	NUT Midline Carcinoma, drug resistance, cancer, and relapse
Date received	19 June 2015
Date approved	28 July 2015

Reference	**2015_DATA_08A**
Project title	Extension to 2015_DATA_08
Institution	Telethon Kids Institute (Australia)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Our group is analysing a number of rare children's diseases to try and find clues to better treatments. This includes paediatric leukaemias, brain tumours, and rarer forms of cancer such as sarcomas and carcinomas. To do this we are using state of the art DNA sequencing technologies, however we are limited in our ability to analyse the data because of difficulties with obtaining control samples from every patient. In order to validate our results, we need to compare our samples to samples from a cohort of normal individuals (i.e., people that do not have cancer or other known diseases). The purpose of this application is to request these normal samples from the EGA study

	EGAS00001000971 which is a UK study that has studied a large number of normal individuals.
Keywords	Rare paediatric cancer
Date received	22 June 2015
Date approved	17 September 2015

Reference	**2015_DATA_09**
Project title	Exome Sequencing in Sudden Infant Death Syndrome
Institution	King's College London (KCL)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Sudden Infant Death Syndrome (SIDS) or 'cot death' is the sudden death of an infant under 12-months old which remains unexplained after thorough testing. It is a leading cause of infant death in the developed world, affecting 1 in 2,500 live births in the UK. Previous research on small numbers of SIDS cases has found heart-related genes being affected in up to 15% of cases. This project aims to assess the role of inherited heart conditions in SIDS by studying genetic samples from over 400 SIDS cases, the largest group ever brought together. Exome sequencing will be used to identify known and new genes that may be responsible for SIDS. The study will determine the role of inherited risk in SIDS and form the basis of guidelines for management of SIDS. This will aid earlier identification of inherited heart conditions and aim to prevent further sudden deaths in the family.
Keywords	Exome sequencing, sudden infant death syndrome, and molecular autopsy
Date received	30 June 2015
Date approved	28 July 2015

Reference	**2015_DATA_10**
Project title	The genetics of obesity
Institution	Wellcome Sanger Institute
Application type	Data access / Genetics
Cohort	NCDS
Abstract	We are interested in how genes influence human body weight. We are currently looking at severely obese individuals and comparing them to non-obese controls to find genetic variants that increase one's risk of being obese. We are currently using controls that are not obese but are diseased. The 1958 exome-sequencing data will enable us to check variants we find are specific to obesity and not due to the disease in controls. We will also be able to explore in more detail the effect of variants on BMI in the population using the exome-sequence and exome-chip genotype data available in the 1958 samples.
Keywords	Obesity and genetics
Date received	6 July 2015
Date approved	28 July 2015

Reference	2015_DATA_11
Project title	Exome Sequencing in Familial Amyotrophic Lateral Sclerosis
Institution	King's College London (KCL)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Amyotrophic lateral sclerosis is a disease of the brain, nerves, and muscles, with an average age of onset of about 50. The average life expectancy is under 3 years from first symptoms, normally due to respiratory failure. We are attempting to determine the genetic causes of this disease, especially amongst patients with affected family members, by sequencing the portion of their genomes that encode for proteins. In order to determine which genetic variants are likely to be involved in disease we need access to a similar dataset from people without the condition.
Keywords	ALS, amyotrophic lateral sclerosis, and whole exome sequencing
Date received	30 July 2015
Date approved	16 December 2016

Reference	2015_DATA_12
Project title	Secular change in the impact of mental disorder risk alleles on psychosocial development.
Institution	Cardiff University
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Our analyses of UK cohort data suggest that children with mental health problems today face greater social difficulties, educational problems, and a poorer mental health prognosis than in previous generations. One unanswered question is whether this just reflects changes in the way that parents fill in mental health symptom questionnaires. We plan to repeat our analyses using a directly equivalent starting point. In particular, advances in psychiatric genetics have allowed researchers to estimate individual genetic risk scores for psychiatric problems. We will use these risk scores to examine whether the prognosis and social and educational impact for those at high genetic risk has also increased over time. To do this we will compare social, educational, and mental health outcomes for children at high vs low genetic risk in two UK birth cohorts (NCDS born 1958 and ALSPAC born in the early 1990s).
Keywords	Childhood, trends, neurodevelopmental, genetic risk score, and polygenic
Date received	3 August 2015
Date approved	17 September 2015

Reference	2015_DATA_13
Project title	Identifying the impact of mental disorder risk alleles on childhood neurodevelopment

Institution	Cardiff University
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Mental disorders are common and contribute more to the global burden of disease than any other type of disorder. Most mental disorders are thought to have their origins at least in part in childhood, although the characteristics in childhood may not be the same as the adult disorder. Mental disorders are heritable, and studies have started to identify specific genetic variants that are associated with these. Genome-wide association studies involving large numbers of patients and controls can be used to provide an estimate of individuals' genetic risk profiles for different mental disorders. We aim to identify childhood neurodevelopmental precursors in the general population that are associated with the genetic risk for adult mental disorders. We will then track longitudinally these risk effects using additional data on mental health and functional outcomes (ill health and pervasive psychosocial difficulties) collected later in development. Our main objective is to identify early origins of mental disorders in order to inform programmes that aim to prevent and target mental disorders as early as possible.
Keywords	Childhood, neurodevelopmental, genetic risk score, polygenic, and bullying
Date received	10 August 2015
Date approved	17 September 2015

Reference	**2015_DATA_14_EXT**
Project title	Extension of data approval to investigate genetic copy number variation on common complex anthropometric traits, namely height, body mass index (BMI), weight, and waist-hip ratio
Institution	University of Leicester
Application type	Data access / Genetics
Cohort	NCDS
Abstract	As well as changes in the sequence of our DNA, genetic variation between individuals can occur as changes in the number of copies we have of certain stretches of DNA, for example, genes. Whilst genotyping array experiments primarily capture sequence variation, it is possible to use the raw data from these experiments to measure "copy number variation" (CNV). In this project, we will utilise this raw data to measure CNV in British 1958 birth cohort participants. We will test for association of CNV with the following related traits: height, weight, body mass index (BMI), and waist-hip ratio (WHR). The results from our analysis will be combined with results from other cohorts in the GIANT consortium in a single large experiment. BMI is a risk factor for multiple chronic diseases and of major public health importance. The results of this will provide important insights into the biological processes underlying height, weight, BMI, and WHR.
Keywords	Copy number, height, weight, waist-hip, BMI, and GiANT consortium

Date received	10 August 2015
Date approved	17 September 2015

Reference	**2015_DATA_15**
Project title	Corticospinal motor neuron development control genes as candidates for human ALS susceptibility.
Institution	King's College London (KCL)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	A key problem in genetics is the interpretation of genetic variation and whether it is relevant to disease. This is particularly an issue for amyotrophic lateral sclerosis (ALS). Current evidence suggests some of ALS is caused by a few genetic variants common in the general population, such as the hexanucleotide expansion in the gene C9orf72, but most is caused by variants much rarer in the population. We have chosen to study developmental genes that control growth of corticospinal motor neurons (CSMN; “upper motor neuron”). There is evidence that rare variants in these genes are detrimental to health of CSMNs. We have analysed CSMN developmental genes using multiple methods. Through a combination of analyses including DNA-sequence analyses, exomic analyses, and gene expression we have found evidence that variation in three of our candidate genes increases the risk of ALS.
Keywords	Amyotrophic lateral sclerosis, rare variants, gene-level burden analysis, DNA sequence analyses, candidate gene, exome, gene expression, and methylation
Date received	20 August 2015
Date approved	16 December 2015

Reference	**2015_DATA_16**
Project title	Contribution of rare germline variants to risk of male breast cancer in the UK population
Institution	Institute of Cancer Research (ICR)
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Although predominantly a disease that occurs in women, breast cancer also affects men – there are on average 350 new cases of breast cancer in men every year in the UK. Because it is quite uncommon, little is known about why men get breast cancer. However, it now appears that some risk factors, particularly those that are inherited (genetic risk factors) are common to both male and female forms of the disease. For a subset of these genetic factors, the risk of breast cancer attributable seems to be larger in men than in women. This suggests that analysis of a modest number of cases of male breast cancer may uncover additional novel risk factors for the disease. We are currently

	conducting such a study by sequencing the DNA of more than 1,000 men with breast cancer.
Keywords	Breast cancer, male breast cancer, DNA sequencing, and Germline predisposition
Date received	11 November 2015
Date approved	18 March 2016

Reference	**2015_DATA_17**
Project title	Identifying the impact of alcohol dependence on family formation and dissolution: from inherited propensities to contextual vulnerabilities
Institution	University of Oxford
Application type	Data access / Genetics
Cohort	NCDS
Abstract	Alcohol dependence (AD) is a complex condition influenced by both genetic and socio-environmental factors, with a lifetime prevalence of 12.5%. Previous studies have shown that AD has a strong genetic component. The goal of this project is to examine the relationship between AD and demographic life course events of family formation (postponement of childbirth, union formation) and dissolution (divorce). We propose to use the 1958 British birth cohort to investigate how genetic propensities for AD and social influences and gene-environmental interactions impact family formation and dissolution patterns. This work is important to not only help us understand the relationship between AD and family events, but also examine causal relationships and the importance of genetic and social determinants in shaping family life course trajectories.
Keywords	Alcohol consumption, fertility behaviour, and family formation
Date received	16 November 2015
Date approved	16 December 2015

Reference	**2015_DNA_01**
Project title	Clinical relevance of cerebral microbleeds in stroke (CROMIS-2)
Institution	National Hospital for Neurology and Neurosurgery (NHNN)
Application type	Samples access / DNA
Cohort	NCDS
Abstract	Previous studies confirm that genetic risk factors play a substantial role in ICH risk and also provide evidence for a role of inherited genetic variation in ICH haematoma volume and outcome. Previously identified risk variants account for a limited proportion of identified genetic influence on ICH risk, and multiple additional loci likely remain unidentified. In a large international GWAS on ICH, with samples from the USA, UK, Europe, and Asia, a discovery phase and replication will be carried out on approximately 20,000 ICH and matched controls. The UK and CROMIS2 will contribute several thousand ICH cases with matched

	Wellcome controls from the 1958 cohort to this ongoing large GWAS study run from Boston looking to identify novel ICH genes.
Keywords	Intracranial haemorrhage
Date received	16 November 2015
Date approved	28 July 2015

Reference	**2015_DNA_02**
Project title	Sequencing projects of rheumatoid arthritis
Institution	Brigham and Women's Hospital (United States)
Application type	Samples access / DNA
Cohort	NCDS
Abstract	Sequencing DNA samples from patients with rheumatoid arthritis and healthy controls to identify rare variants or genes and gene sets enriched for rare variants associated with rheumatoid arthritis. A total of 1,080 genes related with rheumatoid arthritis or other immune traits are selected as targets of sequencing. About 2,000 cases and 1,300 controls will be sequenced. The samples from 1958BC will be used for controls.
Keywords	Rheumatoid arthritis, genetics, and rare variants
Date received	19 November 2015
Date approved	30 September 2015

Reference	**2015_DNA_03**
Project title	DNA methylation profiling in 1958BC samples to assess epigenetic responses to social and environmental cues in early life and over the life course across four UK population-based cohorts
Institution	King's College London (KCL)
Application type	Samples access / DNA
Cohort	NCDS
Abstract	The epigenome provides a mechanism of interaction between the genome with the environment, and we hypothesise that early life stimuli and exposures over the life course leave an epigenetic mark. The proposal aims to explore epigenetic profiles in up to 200 whole blood samples from the 1958 birth cohort, to contribute towards a large discovery analysis of 4,000 samples from four British cohort studies in order to identify epigenetic signatures of early life experience and exposure to social, environmental, and biological stimuli over the life course, linking findings to changes in physical and cognitive function during ageing. Each cohort captures a range of early life experiences, longitudinal health measures, and lifestyle questionnaire data from adult life and has available DNA samples. The proposal will address four research aims, with the overall purpose of identifying epigenetic signatures of early-life events and exposures over life-course that may have impacts on functional health trajectories.
Keywords	DNA methylation and Illumina Infinium HumanMethylationEPIC BeadChip
Date received	20 November 2015

Date approved	01 August 2016