|  |  |
| --- | --- |
| Job Description |  |
| Research Associate/Senior Research Associate – Statistical Genetics, Genetic Epidemiology, Bioinformatics or Data Science(Innovative Medicines Initiative, BigData@Heart Award) |  |
| Department: UCL Institute of Health Informatics | Grade: 7 or 8 (depending on level of experience) |
| Location: 222 Euston Road, London, NW1 2DA | Hours: 1 FTE (5 days per week) |
| Duration: Available immediately until 31 October 2021 in the first instance and may be extended subject to funding. |  |

#### Reports to:

**LONDON’S GLOBAL UNIVERSITY**

Professor Aroon Hingorani, Professor of Genetic Epidemiology, Institute of Cardiovascular Science

Dr Tom Lumbers, Institute of Health Informatics, UCL; Medical and Population Genetics, Broad Institute of MIT and Harvard

#### Context

**Who we are** We are seeking to expand our team of scientists working at the interface of genomic data and phenomic information from electronic health records (EHR). Broadly we seek to exploit a wide range of genomic resources, including those linked to EHR resources, to better understand the genetic architecture of cardiac failure. The resources available include the 100,000 Genomes Project, UK Biobank and the HERMES Consortium. Our team at UCL plays leading roles in the Genomics England Clinical Intepretation Partnership (GeCIP) for Electronic Health Records, Health Data Research UK (HDR UK), in multi-omic cohorts such as UCLEB Consortium, in Biomedical Research Centre initiatives to embed genomic information in clinical care, and in international disease-focused consortia such as HERMES. In addition, the UCL Centre for Translational Genomics has developed Drug Target Database (DTAdB), a data resource that assimilates participant level or public domain summary level genetic association data with mRNA expression, proteomics and metabolomics measures, imaging and physiological traits and disease end points to facilitate gene mapping, drug target validation analyses and multi-trait genetic analysis.

**Scientific problem** The particular focus of this post is to generate new understanding of the genetic architecture of cardiac failure, including sub-types and progression phenotypes, through the analysis of large-scale genomic resources linked to EHR.

**Who you are** You will have a PhD in statistical genetics, genetic epidemiology or bioinformatics with a relevant publication track record. Experience in the analysis of sequence data and the application of complex statistical models would be advantageous. The postholder will be motivated and resourceful, an effective independent worker, and a collaborative team member. We place a strong emphasis on supporting career development and on the potential for independence. This post is funded through the Innovative Medicines Initiative Big Data@ Heart (<https://www.bigdata-heart.eu/>). Suitable candidates will be supported in applications for Fellowships. You will be jointly supervised by Professor Aroon Hingorani, Professor Harry Hemingway and Dr Tom Lumbers.

#### Main purpose of the job

The postholder will be expected to plan and execute analysis to progress understanding of the genetic architecture of heart failure. The postholder will work to discover new genes and pathways for cardiac failure using a range of approaches, including working across common and rare variation and testing for interaction between polygenic risk and variants implicated in Mendelian disease.

#### Key Working Relationships at UCL

Professor Aroon Hingorani (UCL Institute of Cardiovascular Science)

Dr Tom Lumbers (HERMES Consortium, UCL Institute of Health Informatics, Broad Institute of MIT and Harvard)

Professor Harry Hemingway (UCL Institute of Health Informatics, Health Data Research UK London)

Dr Karoline Kuchenbaecker (Statistical Genetics, UCL Genetics Institute)

Dr Spiros Denaxas (Structured EHRs, CALIBER, UCL Institute of Health Infomatics)

Dr Chris Finan (Bioinformatics, UCL Insitute of Cardiovasular Science)

Dr Stephanie Mueller (Statistical Genetics, UCL Insstitute of Health Informatics, EHR GeCIP)

#### Key Working Relationships at EHR GeCIP

Professor Richard Dobson (EHR GeCIP, unstructured and structured EHRs, King’s College London and UCL Institute of Health Infomatics)

Dr Claudia Langenberg (EHR GeCIP, University of Cambridge)

#### Duties and responsibilities:

***Research***

* Analyse CALIBER to develop and validate EHR phenotypes relevant to cardiac dysfunction and heart failure.
* Analyse phenotype of cardiac dysfunction in participants from the 100,000 Genomes Project and UK Biobank to assess the contribution of common and rare variation to cardiac failure.
* Analyse and enrich data from the HERMES heart failure GWAS Consortium (>40 studies, including: >8 national biobanks, >20 clinical trials, 8 large pharmaceutical companies).
* Play a leading role in writing papers reporting findings arising from this work.
* Contribute as author on a number of high impact peer reviewed research publications in general readership clinical or science journals, as well as genetics venues.
* Prepare work in progress presentations for weekly in-house meetings, and national and international conferences
* Develop a portfolio of research to help with personal and career development.
* The postholder will generate efficient and sharable code which can be used in different datasets and methods which are transferable to other diseases.

In pursuit of these research responsibilities the postholder will work well in an interdisciplinary team and be self-starter, taking the initiative to develop and lead research projects based initially within 100,000 Genomes Project, UK Biobank, HERMES and BigData@Heart, with the support of the wider team.

* Submit research publications to refereed journals on a regular basis and present findings at national and international conferences.
* Support ongoing research projects within the Institute, providing supervision and advice to other team members. Work closely with colleagues at other Universities in HERMES, HDR UK and in the EHR GeCIP to deliver the programme of research.
* Contribute to the development of grant applications and manuscripts in collaboration with others in the Institute.

***Teaching and supervision***

Subject to the regulations of the funding bodies, appropriately qualified members of research staff will be expected to contribute to teaching or other departmental activities. At the present time, this is estimated to amount to an absolute maximum of 30 contact teaching hours per annum for full-time staff, with opportunities to develop and deliver teaching on our Master’s programmes (in Health Informatics, in Health Data Science and in Health Data Anlalytics), our Short Course programme, as well as provising informal advice and training of other research staff and students as appropriate.

***Personal development opportunities***

We will strongly encourage and support the postholder to develop their research career; this may include personal fellowship applications, where appropriate. The postholder will be expected to formulate a personal development plan and to capitalise on training opportunities at the Institute and/or through UCL’s Graduate School, as appropriate.

#### Background Information

***Institute of Health Informatics (IHI)***

The aim of the Institute of Health Informatics is to conduct high quality research that leverages big data and health and bioinformatics approaches to improve health at local, national and international levels. Our cutting edge programmes of research span the areas of Discovery Science, Precision Medicine, Learning Health Systems, Public Health and Citizen Driven Health. These programmes of research are underpinned by: excellent secure data infrastructure; adding value to key linked data resources to make data research-ready; development of multidisciplinary methodological expertise; engagement with patients, the public and health services; strong local, national and international academic partnerships and excellent multidisciplinary training programmes. The Institute Director is Professor Harry Hemingway.

***Health Data Research UK (HDR UK) London***

Launched in February 2018, HDR UK is the new national institute for data science for health. HDR UK is funded initially with £54m by a partnership of 9 funders including MRC, EPSRC, ESRC, BHF, NIHR, Wellcome Trust and the governement offices of Scotland, Wales and Northern Ireland. HDR UK has four priority areas: actionable analytics including EHR phenotyping and AI, multi-omics and precision medicine, randomised trials and public health. The London Site of HDR UK comprises UCL (co-ordinating), Imperial, King’s College London, London School of Hygiene % Tropical Medicine and Queen Mary University London. The national Director of HDR UK is Professor Andrew Morris and the London Director is Professor Harry Hemingway.

***Data resources at UCL***

The Institute is host to several international genetic consortia relevant to cardiac failure, including HERMES (heart failure; hermesconsortium.org), UCLEB PMID: 23977022), and GENIUS-CHD (http://www.genius-chd.com/). The results from two completed large scale GWAS meta-analyses of heart failure (>40K cases) and heart failure prognosis (>30K cases) from HERMES are available immediately for analysis. In addition, a range of genomic linked EHR resources are available at the Institute. As lead of the EHR GeCIP (Genomics England Clinical Interpretation Partnership) domain, UCL coordinates a programme of work to integrate coded EHR phenotypes with whole genome sequence data for genetic discovery. The Institute is also engaged in a wide range of studies in UK Biobank, which is being enriched for exome sequence. The UCLH BRC Informatics theme has launched the AboutMe programme to embed genomic information in unselected NHS patients. The UCL Centre for Translational Genomics (PI Hingorani) has developed Drug Target Database (DTAdB). This is a bioinformatics pipeline that assimilates participant level or public domain summary level genetic association data with mRNA expression, proteomics and metabolomics measures, imaging and physiological traits and disease end points from multiple sources and maps them to drug targets, associated compounds and effects of drug action in human from clinical trials.gov and the Cochrane database. The anchoring DNA reference sequence, and Crick’s central dogma (the unidirectional flow of information from gene to RNA to protein and beyond) allows data traversal from disease to drug or vice versa. DTAdB is thus capable of using genomic data to map targets to diseases, and to profile effects of target perturbation on transcriptome, proteome, metabolome and other intermediate phenotypes, as well as disease end-points. In addition to the focus on drug development, this resources is available to support the mapping of genetic association signals to candidate genes and to perform secondary cross-trait analysis. The CALIBER resource (linked primary care and secondary care data on about 10m individuals), although it does not contain genetic information, has phenotypic details and diagnosis details of a wide range of single gene disorders, including cardiomyopathies. The BigData@Heart Consortium provides further opportunities through collaboration.

# Person specification

| Criteria | Essential or Desirable (Grade 7) | Essential or Desirable(Grade 8) |
| --- | --- | --- |
| **Qualifications, experience and knowledge** |  |  |
| PhD in statistical genetics, genetic epidemiology, bioinformatics, computational biology (or related discipline) or a similar level of experience | Essential | Essential |
| Experience and advanced skills analysing and interpreting genetic and genomic data | Essential | Essential |
| Experience and advanced skills analysing and interpreting next generation sequence data | Desirable | Essential |
| Experience developing and implementing analysis pipelines within in a UNIX/ Linux environment | Essential | Essential |
| Scripting skills and experience in at least one coding language e.g. Perl/ Python/ BASH | Essential | Essential |
| Advanced skills in at least one statistical software package e.g. R | Essential | Essential |
| Experience a range of bioinformatics tools for gene mapping and cross-trait analysis | Desirable | Essential |
| Experience of manipulating and analysing large, high dimensional datasets including EHR | Desirable | Essential |
| Experience in relational databases and SQL or similar language(s) | Desirable | Essential |
| Experience in designing, implementing and evaluating statistical learning methods in large datasets | Desirable | Essential |
| Proven track record of publishing in peer reviewed journals | Essential | Essential |
| Understanding of information governance, privacy and security issues | Essential | Essential |
| Working in multi-disciplinary teams and international collaborations | Desirable | Essential |
| Presentation of research through conferences or lectures | Essential | Essential |
| **Skills and abilities** |  | Essential |
| Ability to manage own time and prioritise workload to meet project deadlines | Essential | Essential |
| Ability to write up the methods and results of statistical analyses for clinical audiences | Essential | Essential |
| Ability to present orally the methods and results of statistical analyses | Essential | Essential |
| Desire to develop and extend role | Essential | Essential |
| **Personal attributes** |  |  |
| Self-motivated and ability to work on own initiative | Essential | Essential |
| Willingness and ability to work collaboratively with colleagues | Essential | Essential |
| Commitment to high quality academic research | Essential | Essential |
| Commitment to continuing professional development | Essential | Essential |
| Commitment to fostering a positive work/learning environment | Essential | Essential |
|  |  |  |

# Apply

|  |
| --- |
| To apply for this position visit:ucl.ac.uk/jobs |