CANDIDATE BRIEF

Data Scientist, Faculty of Medicine and Health
Leeds Institute of Health Sciences

Salary: Grade 8 (£40,792 – £48,677 per annum)

Reference: MHIHS1231

Closing date: 2nd October 2019

Interview date: 15th October 2019

Fixed term for 12 months.

We will consider job share and flexible working arrangements

Data Scientist
School of Medicine, Leeds Institute of Health Sciences

Are you an enthusiastic and driven individual with an excellent working knowledge of data systems, processing and analysis? Do you want take a leading role in the data processing and analysis for a suite of trials representing some of the largest studies in Lung Cancer Screening active anywhere in the world? Do you want to join a large, successful multi-disciplinary team delivering these studies which have the potential to shape how Lung Cancer Screening and Smoking Cessation together might be delivered to many thousands of future eligible participants?

Leeds Institute of Health Sciences is seeking to appoint an enthusiastic and motivated Data Scientist/Programmer with experience in data management and programming using large datasets from multiple sources in health care.

The post will provide an excellent opportunity to contribute to high impact health research in the field of cancer screening and specifically to aid the conduct of the Yorkshire Lung Screening Trial and its associated studies.

The successful applicant should hold a Master's degree (or have appropriate equivalent experience), have good communications skills, demonstrable programming and analytic skills and experience of working with large, complex datasets. The candidate will be expected to use programming languages such as SQL to collate and manage data, provide oral and written reports, contribute to publications and work within a team environment.

The Yorkshire Lung Screening Trial (YLST) is the UK's largest randomised trial of Lung Cancer Screening, funded by Yorkshire Cancer Research. Lung Cancer Screening is provided for high-risk participants in a community setting as part of a Lung Health Check. Alongside the main trial is a biomarker sub-study whereby participants attending Lung Health Checks separately consent to blood sample collection to investigate possible circulating markers of lung cancer. Embedded within YLST is the Yorkshire Enhanced Stop Smoking Study (YESS) whereby current smokers attending the mobile units are invited to participate in a study of a personalised stop-smoking intervention in the context of screening. YESS is the largest study looking at the best way to integrate smoking cessation into screening anywhere in the world.

The successful applicant will be a member of the YLST team with overall responsibility for Data Management and facilitating trial conduct within YLST and the biomarker sub-study. Specifically, the post holder will use programming and analytic skills to manage and collate a number of relevant healthcare datasets for the purpose of ongoing trial management, clinical processes, interim monitoring and analyses. Important sources of data for the trial include primary care data from participating general practices, secondary care data from the PPM+ system in Leeds Teaching Hospitals, YLST-specific data, and national cancer and death registrations currently under the auspices of NHS digital. . You will be based in Leeds with the core trial team and will also work closely with the PPM+ team and with our external partners, including the YLST Statisticians and YESS team to facilitate data management. The post provides an excellent opportunity to contribute to high impact health research in the field of cancer screening. The successful applicant should hold a masters degree, have good communications skills, demonstrable programming and analytic skills and experience of working with large, complex datasets. The candidate will be expected to use programming languages such as SQL to collate and manage data, provide oral and written reports, contribute to publications and work within a team environment.

The post holder will report to Professor Richard Neal, and be encouraged to develop their own ideas for improving research and study conduct in an area that requires management and interrogation of large complex healthcare datasets.

Funding is available for 12 months in the first instance.

What does the role entail?

- Supervision of data management for the Yorkshire Lung Screening Trial
- Data management and manipulation of large and complex relational databases, including adapting existing methods and developing new approaches.
- Creating and testing programming algorithms using languages such as SQL to check data for errors (e.g. consistency, range, duplication checks), clean data, update and derive data where appropriate, extract, import, merge and export data in a format suitable for analysis.
- Undertake necessary training relevant to data protection and security, good clinical practice and any other necessary training to work in the Leeds Institute of Health Sciences and gain access to NHS records within Leeds Hospital Trust.
- Liaise closely with the Information Team in Leeds Teaching Hospitals to facilitate extraction of data from Electronic Patient Record (PPM+)
- Develop coding strategies and methods for identifying eligible participants/cases/ conditions/ treatment regimes.

- Contribute to the development and maintenance of Standard Operating Procedures for data provision, data manipulation and management, and archiving and destruction of datasets
- Write programs to allow querying and interrogation of the data to facilitate trial conduct and monitoring.
- Running data extraction and randomisation processes for each round of screening from Primary Care databases, according to Standard Operating Procedures
- Work to agreed research project objectives under the direction of the Chief Investigator and Chief Statistician
- Provide advice on data management issues
- Training and supervision of other members of the clinical team regarding data entry and processes
- Contribute to key study documentation including the trial protocol and standard operating procedures
- Scoping exercises for feasibility of proposed related research
- Collaborate with colleagues in partner institutions, and research groups
- Develop own area of research interest
- Collaborate in the preparation of research publications and reports
- Manage own day to day research and administrative activities
- Supervising other members of the trial team with responsibility for data processes
- Arranging permissions for transfer of data to associated institutions for analysis

What will you bring to the role?

As a Data Scientist you will have:
- A relevant Master's degree, or have appropriate equivalent experience
- Previous experience of data management of large scale datasets and relational databases, preferably in health care, and/or other electronic medical records.
- Be able to pay attention to detail and apply analytical thinking to detect errors and provide solutions
- Experience of using programming to check data for errors and clean data, create algorithms for categorising, deriving or updating data and import, merge and export data
- Strong programming skills in relevant languages such as SQL.
- Excellent IT skills and familiarity with MS packages including Excel, Word, and PowerPoint.
- Knowledge of data protection and confidentiality issues
- Excellent interpersonal skills, including effective communications with people from different scientific or clinical disciplines.
- Good written and oral communication skills, including the ability to write reports and study documentation.

- Ability to plan, implement and deliver work, in order to meet deadlines and associated activities, independently and as part of research team
- To be able to demonstrate potential for applying scientific knowledge and skills to improve data management techniques for studies involving large and varied, complex datasets in healthcare.

You may also have:
- Experience of independently managing a discrete area of a research project
- Good knowledge of Information Governance Procedures

How to apply

You can apply for this role online; more guidance can be found on our How to Apply information page. Applications should be submitted by **23.59** (UK time) on the advertised closing date.

Contact information

To explore the post further or for any queries you may have, please contact:

Dr Mat Callister - Tel 0113 206 4159 - matthew.callister@nhs.net

Additional information

Find out more about the Faculty of Medicine and Health.

Find out more about Athena Swan in the Faculty of Medicine and Health.

Working at Leeds
Find out more about the benefits of working at the University and what it is like to live and work in the Leeds area on our Working at Leeds information page.

Candidates with disabilities
Information for candidates with disabilities, impairments or health conditions, including requesting alternative formats, can be found on our Accessibility information page or by getting in touch with us at disclosure@leeds.ac.uk.

Criminal record information

Rehabilitation of Offenders Act 1974

A criminal record check is not required for this position. However, all applicants will be required to declare if they have any 'unspent' criminal offences, including those pending.

Any offer of appointment will be in accordance with our Criminal Records policy. You can find out more about required checks and declarations in our [Criminal Records](#) information page.