### Purpose of the Post

This is an exciting opportunity for an early-career data scientist to join a leading research group to design, build and implement a platform for analysis of longitudinal data generated from the Breathing Together study. Breathing Together is a longitudinal cohort study of 1,000 children who are being followed from birth to age three to see who develops wheeze and asthma and an additional cross-sectional study of 300 preschool to school aged children with severe wheeze or asthma and aged matched controls. In addition to clinical record data, we also have monthly questionnaire data, epithelial cell function and transcriptomics data, microbiome analysis data, blood cell transcriptomics, genetics and urinalysis data, all collected at multiple timepoints. You will be responsible for developing systems to link and analyse these disparate data sources.

You will be based in the Section of Inflammation, Repair and Development based at the South Kensington Campus, and you will work closely with clinicians and researchers to deliver this work. The project is funded by the Wellcome Trust. You will be working on a completely unique data set in a rapidly evolving area of clinical science and a very large potential for clinical impact. You will develop skills in natural language processing, big data science, bioinformatics and data visualisation.

### Key Responsibilities

The key tasks of the post holder will be:

- Data cleansing and partitioning of large datasets
- Working with the Breathing Together consortium to create a secure analysis platform for the scaled analysis of complex health data.
- Predictive and advanced analytics including machine learning and natural language processing
- Contribute to reports for funders, peer-reviewed journals, policymakers, academics, and other stakeholders
## Job Description

### Person Specification

**Requirements**

Candidates/post holders will be expected to demonstrate the following:

<table>
<thead>
<tr>
<th>Qualification</th>
<th>Essential (E)/Desirable (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree or equivalent experience in IT, computer science, information management, bioinformatics or a related numerate or scientific discipline.</td>
<td>E</td>
</tr>
<tr>
<td>Master’s degree in information technology or science subject with a strong quantitative focus</td>
<td>E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experience</th>
<th>Essential (E)/Desirable (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience in programming with data science focused language (R/Python)</td>
<td>E</td>
</tr>
<tr>
<td>Significant software development experience with an emphasis on large-scale scientific data management</td>
<td>E</td>
</tr>
<tr>
<td>Experience with data science modules and libraries (keras, tensorflow, scikit-learn, nltk)</td>
<td>E</td>
</tr>
<tr>
<td>Experience in working with medical data</td>
<td>E</td>
</tr>
<tr>
<td>Experience in handling/processing large and complex datasets.</td>
<td>E</td>
</tr>
<tr>
<td>Well trained in machine learning and/or applied statistics.</td>
<td>E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Knowledge</th>
<th>Essential (E)/Desirable (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proven track record of biomedical research</td>
<td>E</td>
</tr>
<tr>
<td>Knowledge of big data platforms (e.g. tranSMART)</td>
<td>E</td>
</tr>
<tr>
<td>Knowledge of data visualisation platforms</td>
<td>D</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Skills &amp; Abilities</th>
<th>Essential (E)/Desirable (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent verbal and written communication skills.</td>
<td>E</td>
</tr>
<tr>
<td>Good administrative and organisational skills including excellent record keeping.</td>
<td>E</td>
</tr>
<tr>
<td>Ability to complete work to deadlines while maintaining a high degree of accuracy and excellent attention to detail.</td>
<td>E</td>
</tr>
<tr>
<td>Ability to prioritise own workload.</td>
<td>E</td>
</tr>
<tr>
<td>Ability to work both as part of a team and independently, using initiative where required.</td>
<td>E</td>
</tr>
<tr>
<td>Desire to continue with professional development and ability to acquire new skills quickly.</td>
<td>E</td>
</tr>
<tr>
<td>Problem solving skills and enthusiastic approach to new challenges.</td>
<td>E</td>
</tr>
<tr>
<td>Quick learner and self-motivated.</td>
<td>E</td>
</tr>
<tr>
<td>An open and positive attitude to working in a constantly changing environment.</td>
<td>E</td>
</tr>
</tbody>
</table>

Please note that job descriptions cannot be exhaustive and the post-holder may be required to undertake other duties, which are broadly in line with the above key responsibilities.

Imperial College is committed to equality of opportunity and to eliminating discrimination. All employees are expected to follow the 7 Imperial Expectations detailed below:

1) Champion a positive approach to change and opportunity
2) Communicate regularly and effectively within and across teams
3) Consider the thoughts and expectations of others
4) Deliver positive outcomes
5) Encourage inclusive participation and eliminate discrimination
6) Support and develop staff to optimise talent
7) Work in a planned and managed way

Employees are also required to comply with all College policies and regulations paying special attention to:

- Confidentiality
- Conflict of Interest
- Data Protection
- Equal Opportunities
- Financial Regulations
- Health and Safety
- Information Technology
- Smoking
- Private Engagements and Register of Interests
Job Description

They must also undertake specific training and assume responsibility for safety relevant to specific roles, as set out on the College Website Health and Safety Structure and Responsibilities page.

The College is a proud signatory to the San-Francisco Declaration on Research Assessment (DORA), which means that in hiring and promotion decisions, we evaluate applicants on the quality of their work, not the journal impact factor where it is published. For more information, see https://www.imperial.ac.uk/research-and-innovation/about-imperial-research/research-evaluation/

The College believes that the use of animals in research is vital to improve human and animal health and welfare. Animals may only be used in research programmes where their use is shown to be necessary for developing new treatments and making medical advances. Imperial is committed to ensuring that, in cases where this research is deemed essential, all animals in the College’s care are treated with full respect, and that all staff involved with this work show due consideration at every level. http://www.imperial.ac.uk/research-and-innovation/about-imperial-research/research-integrity/animal-research/

Committed to equality and valuing diversity, we are an Athena SWAN Silver Award winner, a Stonewall Diversity Champion, a Disability Confident Employer and work in partnership with GIRES to promote respect for trans people.