PhD studentship

Part-funded by the

NIHR Maudsley Biomedical Research Centre

Bioinformatics

Studentship to commence October 2018

Contents

NIHR Maudsley Biomedical Research Centre (BRC)

NIHR Biomedical Research Centres are funded to support people and/or patient-focused early translational (experimental medicine) research, the aim of which is to translate discoveries from basic/discovery science into clinical research, and through to benefits for patients, the health system and for broader economic gain.

On September 16 2016 the Secretary of State for Health announced that the Department of Health has awarded £66 million funding over the next five years to the National Institute for Health Research (NIHR) Biomedical Research Centre (BRC) at South London and Maudsley NHS Foundation Trust and the Institute of Psychiatry, Psychology & Neuroscience at King's College London.

The award represents a substantial uplift in funding compared to the previous BRC funding round, and demonstrates the government's continued commitment to the current NIHR Maudsley BRC, allowing the research centre both to build on its current work and expand into new areas including substance use, obesity, pain and mobile health technology.

The expanded NIHR Maudsley BRC will bring together scientists, clinicians, mental health professionals, service users and carers, to improve clinical care and services across the field of mental health. The investment in the NIHR Maudsley BRC will allow research into ground-breaking treatments and care for mental health and dementia.

NIHR Maudsley BRC Strategy

There are four major elements to the NIHR Maudsley BRC strategy for the coming 5 years, reflected in aims of the 17 themes:

- **Precision psychiatry**: Bringing together insights from cognition, behaviour, genomics and brain imaging, we will develop biologically-informed strata of psychiatric syndromes, with the ambition to develop and provide more individually tailored treatment
- **Novel therapeutics**: Using the access to our large databases, electronic consent for contact procedures, and our dedicated experimental medicine Clinical Research Facility (CRF), we will undertake trials of new pharmacological, neuromodulation and psychological treatments
- **Translational informatics**: By using our bespoke natural language processing algorithms and 'smart agents', we will use informatics to influence treatment choice, increase adherence, improve health behaviours and increase patient empowerment, all of which will benefit patient outcomes and service delivery
- **Mental/physical interface**: We will decrease the 15 years of life lost to serious mental illness by using informatics to identify, prioritise and track the treatment of those with comorbid mental and physical disorders

Clinical disorder focused research themes

Seven clinical disorder focused research themes cover mental health and dementia from cradle to grave:

- Affective Disorders and Interface with Medicine
- Child and Neurodevelopmental Disorders
- Dementia and Related Disorders
- Lifestyle Substance Use & Harms (Substance Use)
- Obesity, Lifestyle and Learning from Extreme Populations (Obesity)
- Pain and headache
- Psychosis and Neuropsychiatry

Technology and methodology focused research themes

Seven technology and methodology focused research themes develop and deploy new approaches to clinical problems:

- Bioinformatics and Statistics
- Biomarkers and Genomics
- Clinical and Population Informatics
- Mobile Health
- Neuroimaging
- Patient and Carer Involvement and Engagement
- Translational Therapeutics

Cross cutting themes

Three cross cutting themes provide enabling infrastructure:

- BioResource
- Clinical Research Facility
- Training and Capacity Development

Bioinformatics and Statistics

Lead: Professor Andrew Pickles

This theme provides the computing infrastructure and expertise in statistics and bioinformatics required to integrate and use the complex multimodal data we have access to, particularly via our CRIS system, biomarkers, and data derived from our Mobile Health theme. We use advanced computer science approaches to improve patient care by developing a “self learning” healthcare system.

Aims

1. Deliver a “panor-omic” view of each patient through integration of –omics data with data derived from patient reports, electronic health records, exposures, social graphs, imaging and other emerging technologies, and to systematically exploit these data to explore precision psychiatry and the mental-physical health interface
2. Develop methodology and implement in-service designs for enhanced learning to optimise treatment selection and combination; where possible, to enable formal learning from routine practice, a key element in delivering a “self-learning” health care system
3. Deploy a programme of translational informatics Applied Intelligence software agents into routine practice to identify and communicate with patients and clinicians, and to monitor and evaluate the effects of novel therapies and changes in clinical practice, resulting in treatments and services that are cheaper, faster, more reliable or less intrusive than traditional methods

Institute of Psychiatry, Psychology and Neuroscience

The Institute is organised into three academic divisions, each comprised of a number of cognate departments. Each Division includes academics and researchers from diverse scientific disciplines, working closely with colleagues across the faculty and our national and international partners:

- **Division of Academic Psychiatry** comprises 6 departments: Addictions Sciences; Forensic & Neurodevelopmental Science; Child & Adolescent Psychiatry; Old Age Psychiatry; Psychological Medicine and Psychosis Studies (https://www.kcl.ac.uk/ioppn/divisions/academic-psychiatry/index.aspx)
- **Division of Psychology & Systems Science** comprises 4 departments: Biostatistics & Health Informatics; Health Service & Populations Research; Social Genetic & Developmental Psychiatry; Psychology; (https://www.kcl.ac.uk/ioppn/divisions/psychology/index.aspx)
- **Division of Neuroscience** comprises 4 departments: Basic & Clinical Neuroscience; Neuroimaging; Developmental Neurobiology; Wolfson Centre for Age-related Diseases (https://www.kcl.ac.uk/ioppn/divisions/neuroscience/index.aspx)

Successful applicants for this studentship will be registered for their MPhil/PhD with King's College London and will be based in the Biostatistics and Health Informatics department at the Institute of Psychiatry, Psychology and Neuroscience (IoPPN).

Please note: The final project details are to be agreed after successful interview.

Project

Important: When applying for this studentship, in addition to the personal statement, please upload a separate single-side A4 document headed with the reference number and title of the project in this catalogue. Please include a statement explaining why you have chosen this project and why you would like to take this forward as a PhD (**maximum 300 words**).

If you wish to apply for one or more of the other studentships we are currently advertising, please upload a *separate A4 sheet for each studentship* you are applying for, stating your preferred project choices from those advertised with the studentship, and a statement about your first choice project (see above). Please ensure each sheet clearly indicates which studentship you are applying for and lists only projects advertised for that particular studentship.

If you wish to discuss a project before you apply, you will find supervisors' names and their contact details listed with each project in the catalogues for each studentship.

Further information about project supervisors can be viewed in the King's College London Research Portal. Under **Researchers**, type the name of the person you wish to view information about.

Please note: The final project details are to be agreed after successful interview.

BIOS-2.01 Modern NLP Applied to electronic health records for stratification and clinical decision making in psychiatry

Primary Supervisor: Professor Richard Dobson

Academic Department: Biostatistics and Health Informatics

Email: richard.j.dobson@kcl.ac.uk

Website: https://kclpure.kcl.ac.uk/portal/richard.j.dobson.html and https://phidatalab.org/

Second Supervisor: Dr Zina Ibrahim

Academic Department: Biostatistics and Health Informatics

Email: zina.ibrahim@kcl.ac.uk

Website: https://kclpure.kcl.ac.uk/portal/zina.ibrahim.html and https://www.zibrahim.info/

Project Description

Background: Mental health faces many challenges in delivering the promise of precision medicine and in some cases is lagging behind some areas of physical health[1,2]. Studies show that the large volume of data stored within Electronic Health Records (EHRs) can provide valuable insight and high-quality indicators of recommended treatments[3-4].

In this project, we want to explore the potential for algorithmic matching of a patient's EHR longitudinal profile, including physical health comorbidities, with similar patients in order to take us a step closer to the goal of being able to deliver more precise care. However, the EHRs are unstructured, heterogeneous, episodic and sparse, necessitating the creation of scalable, efficient, and interpretable computational models capable of automatically learning patient similarity.

Novelty and Importance: Automating the stratification of patients based on similarity between their EHR-mined digital phenotypes to identify likely treatment responses or prognosis for patient subgroups will allow for enormous benefits in terms of better outcomes and adversity minimization.

Continued on next page

[1] Cuthbert BN, Insel TR (2013) Toward the future of psychiatric diagnosis: the seven pillars of RDoC. BMC Med 11:126.
[2] Insel TR The NIMH Research Domain Criteria (RDoC) Project: precision medicine for psychiatry. - PubMed - NCBI. https://www.ncbi.nlm.nih.gov/pubmed/24687194.
[3] Rossom RC, Simon GE, Beck A, et al (2016) Facilitating Action for Suicide Prevention by Learning Health Care Systems. Psychiatr Serv 67:830–832
[4] Silow-Carroll S, Edwards JN, Rodin D (2012) Using electronic health records to improve quality and efficiency: the experiences of leading hospitals. Issue Brief 17:1–40

BIOS-2.01 Modern NLP Applied to electronic health records for stratification and clinical decision making in psychiatry

Primary aim(s): This project aims to:

a) Develop a suite of efficient and scalable deep learning models for automatically measuring similarity between patient longitudinal profiles mined from their unstructured mental and physical EHRs.

b) Build a proof-of-concept decision support tool to evaluate the models' effectiveness in making personalized recommendations to carers.

   a. Evaluate the similarity measures and decision support prototype to predict outcome and drug-related adversity on patients with use cases that could include: a) rheumatoid arthritis with comorbid depression, b) anxiety

Planned research methods and training provided:

1. Pattern recognition module at the Department of Informatics (DoI).
2. Machine learning module at the DoI.
3. Introduction to Natural Language Processing at the Department of Biostatistics & Health Informatics.

Objectives / project plan:

Year 1: Obtain SLaM and KHP research passports. Conduct literature review. Complete the four taught modules. Mine CRIS and KCH data.

Year 2: Develop and evaluate computational models with evaluation on the two use cases

Year 3: Build decision support prototype and further evaluate models.

Two representative publications from supervisors:

1: Wu, H., Toti, G., Morley, K. I., Ibrahim, Z. M., Folarin, A., Jackson, R., ... & Gorrell, G. (2018). SemEHR: A general-purpose semantic search system to surface semantic data from clinical notes for tailored care, trial recruitment, and clinical research. *Journal of the American Medical Informatics Association*.

2: C. Ormandy, Z Ibrahim, R Dobson (2017). Learning Patient Similarity Using Joint Distributed Embeddings of Treatment and Diagnoses. Knowledge Discovery from Healthcare Data, Co-located with the International Joint Conference on Artificial Intelligence (KDH@IJCAI) 1891, 30-35.

Keywords: Health informatics; Deep learning; Text mining; Natural Language Processing; NLP; Learning Health Systems;

BRC Theme/s: Affective Disorders and Interface with Medicine
Child and Neurodevelopmental Disorders
Bioinformatics and Statistics