PhD studentship

Funded by the

NIHR Maudsley Biomedical Research Centre

Project Catalogue

Bioinformatics and Statistics

Studentship to commence October 2018

Contents

Introduction

Welcome to the National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre (BRC) project catalogue for potential candidates wishing to commence a PhD in October 2018 – we hope you will find a project which interests you.

The Maudsley BRC is a collaboration between the Institute of Psychiatry, Psychology and Neuroscience (IoPPN), King's College London – the largest collection of researchers in Europe investigating mental disorders, and the South London and Maudsley NHS Foundation Trust – a leading mental health trust with a long tradition in joining clinical and academic excellence. Most Maudsley BRC researchers, staff and students are based at the IoPPN at the King's College London Denmark Hill campus which is adjacent to the Maudsley Hospital. Within this setting we offer the opportunity to join a thriving group of interdisciplinary researchers with internationally recognised supervisors and we ensure our students benefit from an understanding of the context of their research, producing scientists with a strong translational ethos.

The Maudsley BRC is dedicated to developing better treatments for people with mental and neurological disorders, which collectively cause most of the disease burden in Western societies. Within the BRC we offer projects which are clinically relevant and attempt to bring new innovation to help treat people with mental disorders, dementia and other neurological conditions. This is the most exciting field in biomedical science, the least researched, the most important. And we offer an opportunity to gain research training in a vibrant and exciting centre where doctoral students are highly valued members of our team.

We hope we can look forward to receiving your application.

Professor Matthew Hotopf
Director
Maudsley Biomedical Research Centre

Professor Richard Brown
Training Lead
Maudsley Biomedical Research Centre

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 same department as their first supervisor at the Institute of Psychiatry, Psychology and Neuroscience (IoPPN).

Please note: The final choice of project and project details are agreed after successful interview.

Projects

When applying for the NIHR Maudsley Biomedical Research Centre PhD studentship in the **Bioinformatics and Statistics** theme, please ensure you state your two preferred PhD projects from those listed in this catalogue only**. These should be listed in order of preference and include the number that is assigned to the project and the project title.

For example:

1. BSTA-2.03 Towards personalised medicine for antidepressant drugs: a machine learning approach
2. BSTA-2.01 Understanding transdiagnostic mechanisms in cognitive behavioural interventions for patients with Persistent Physical Symptoms

****Important:** With your application, in addition to the personal statement, please upload a separate single-side A4 document listing your first and second choice projects with a statement explaining why you have chosen your **first choice** project and why you would like to take this forward as a PhD (**maximum 300 words**). If you wish to apply for more than one studentship, a separate A4 document should be uploaded for each studentship making it clear which studentship and which projects you are interested in.

****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 this catalogue.

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 choice of project and project details are agreed after successful interview.

BSTA-2.01 Understanding transdiagnostic mechanisms in cognitive behavioural interventions for patients with Persistent Physical Symptoms

Primary Supervisor: Professor Sabine Landau

Academic Department: Biostatistics and Health Informatics

Email: Sabine.landau@kcl.ac.uk

Website: https://kclpure.kcl.ac.uk/portal/sabine.landau.html

Second Supervisor: Professor Trudie Chalder

Academic Department: Psychological Medicine

Email: Trudie.chalder@kcl.ac.uk

Website: https://kclpure.kcl.ac.uk/portal/en/persons/trudie-chalder.htm

Project Description

Background: Persistent Physical Symptoms (PPS) are associated with profound disability and high health care costs. There is an accumulating body of evidence demonstrating that cognitive behavioural interventions can reduce levels of symptoms and improve functioning in a range of PPS. While it is standard clinical practice to adapt psychological therapies to the patient population it is not clear which components of the cognitive and behavioural interventions are transdiagnostic, that is, address responses shared by patients across the PPS spectrum, and which are disorder specific. Identifying such mechanisms can help clinicians target core mechanisms and develop new psychological interventions for other patient groups with PPS. This PhD project will develop and apply methods for modelling the impact of the symptoms on underlying mechanisms and will provide an opportunity to develop knowledge of psychological theory and interventions and skills in biostatistics.

Novelty and Importance: The project will develop novel statistical modelling approaches so that existing data sources can be fully exploited to develop theories regarding transdiagnostic mechanisms. The methodology will be applied to cognitive behavioural interventions in PPS but the new methods will be relevant to the development of psychological therapies more widely

Primary aim(s):
To identify components that represent transdiagnostic mechanisms of the cognitive and behavioural interventions in order to better target psychological interventions for patient groups with PPS.

Planned research methods and training provided: The student will have access research methods training provided by the BRC and also to the programme of short courses on advanced statistical methodological provided by the Department of Biostatistics & Health Informatics. This will include training in structural equation modelling. In addition, the successful candidate will be able attend external training if required.

Continued on next page

BSTA-2.01 Understanding transdiagnostic mechanisms in cognitive behavioural interventions for patients with Persistent Physical Symptoms

Objectives / project plan:

Year 1: The student will carry out a systematic review of mechanistic theories in PPS and existing methods for assessing transdiagnostic mechanisms and will prepare individual participant data (IPD) from a number of CBT trials for pooling; likely to include the PACE trial for CFS (White et al 2011) and PRINCE Secondary trials for PPS in general. A number of putative mediators and effect modifiers have been measured across trials. This project will develop modelling techniques to assess whether mechanisms are shared across disorders or operate differentially.

Years 2 and 3: We envisage structural equation modelling and IPD meta-analysis/integrative data analysis techniques to play an important role in this methodological project.

Two representative publications from supervisors:

1: Dunn G., Emsley E., Liu H., **Landau S.**, Green J., White I. & Pickles A. (2015) Evaluation and validation of social and psychological markers in randomised trials of complex interventions in mental health: a methodological research programme. *Health Technology Assessment* **19** (93) http://dx.doi.org/10.3310/hta19930

2: Chalder T., Goldsmith KA, White PD, Sharpe M, Pickles AR. (2015) Rehabilitative therapies for chronic fatigue syndrome: a secondary mediation analysis of the PACE trial. *Lancet Psychiatry* **2** (2):141-52. doi.org/10.1016/S2215-0366(14)00069-8

Keywords: Persistent Physical Symptoms; Cognitive Behavioural Therapy; CBT; Transdiagnostic; Mechanisms; Trials;

BRC Theme/s:
Affective Disorders and Interface with Medicine
Bioinformatics & Statistics
Clinical & Population Informatics

BSTA-2.02 Applications of Regularized Structural Equation Modeling to Psychiatric Genetics and Psychometrics

Primary Supervisor: Dr Daniel Stahl

Academic Department: Biostatistics and Health Informatics

Email: daniel.r.stahl@kcl.ac.uk

Website: https://kclpure.kcl.ac.uk/portal/daniel.r.stahl.html

Second Supervisor: Dr Cedric Ginestet

Academic Department: Biostatistics and Health Informatics

Email: cedric.ginestet@kcl.ac.uk

Website: https://kclpure.kcl.ac.uk/portal/cedric.ginestet.html

Third Co-supervisor: Dr Silia Vitoratou

Academic Department: Biostatistics and Health Informatics

Email: silia.vitoratou@kcl.ac.uk

Website: https://kclpure.kcl.ac.uk/portal/silia.vitoratou.html

Project Description

Background: One of the central challenges of modern psychiatry is to identify subsets of patients that are responsive to specific forms of treatment. Such an approach often proceeds by identifying biomarkers for certain diagnostic categories. These biomarkers can be latent, in the sense that they are not directly measured, as in the case of polygenic risk scores, or when developing psychometric instruments.

The characterization of such latent constructs can be performed using a well-established modelling framework, called structural equation modeling (SEM). SEM distinguishes two latent measurement models: reflective and formative. In reflective SEM (e.g. factor analyses), it is assumed that the latent variables influence the response on the observable variables, or items. This is the standard approach in the development of measurement scales in psychometrics. In formative models, by contrast, the observed variables are assumed to be one of the causes of the latent variables. Single nucleotide polymorphisms, for instance, are assumed to influence the latent risk for a particular condition.

Novelty and Importance: Standard SEMs do not perform well when the number of variables is large relative to sample size. Therefore, the routine collection of genetic, proteomic and other –omic data sets; poses fresh challenges to the application of SEM methods for the detection and validation of psychiatric biomarkers. One possible solution to these issues is to combine machine learning methods with SEM into so-called regularized SEMs. Recent methodological advances have provided a set of tools for using regularized SEM in standard statistical software (Jacobucci et al., 2015), but have not yet been applied to psychiatric research.

Primary aim(s): This PhD project will aim to apply the newly introduced regularized SEM methodology to two existing data sets in order to assess its usefulness in comparison to traditional methods.

BSTA-2.02 Applications of Regularized Structural Equation Modeling to Psychiatric Genetics and Psychometrics

Planned research methods and training provided: The student will attend courses covering the use of statistical softwares, R and MPlus, machine learning methods, SEM, as well as psychometrics in KCL and other institutions. Training on the use of genetic data sets will also be provided in the IoPPN.

Objectives/project plan:

Year 1: Familiarization with the basic principles of regularization and SEM, as well as with the R toolbox regSEM to implement regularized SEM.

Year 2: Use regularized SEM for the analysis of the RADIANT depression study, which contains both extensive genomic and clinical data, with special focus on the computation of polygenic risk scores.

Year 3: Apply regularized SEM to the analysis of a large psychometrics data set that has been used for the screening and differential diagnosis of personality characteristics relying on the Temperament and Character Inventory (Cloninger et al, 1994).

References:

Jacobucci, R., Grimm, K.J., and McArdle, J.J. (2016) Regularized Structural Equation Modeling. *Structural Equation Modelling: A Multidisciplinary Journal, 23(4),* 555-566.

Cloninger CR, Przybeck TR, Svrakic DM, Wetzel RD. The Temperament and Character Inventory (TCI). A Guide to its development and use. St. Louis: Center for Psychobiology of Personality, Washington University; 1994.

Two representative publications from supervisors:

1: Iniesta R, Stahl D, McGuffin P. (2016). Machine learning, statistical learning and the future of biological research in psychiatry. Psychol Med. 46(12):2455-65. doi: 10.1017/S0033291716001367

2: Vitoratou, S, Ntzoufras, I., Theleritis, C., Smyrnis, N., Stefanis, N.C. (2015). Temperament and character dimensions assessed in general population, in individuals with psychoactive substance dependence and in young male conscripts European Psychiatry, 30(4) , 474 – 479. doi.org/10.1016/j.eurpsy.2015.01.007

Keywords: Machine learning; Structural equation modelling; Polygenic risk scores; Psychometrics; Personalised medicine;;

BRC Theme/s: Bioinformatics and Statistics
Biomarkers and Genomics
This project links in with most of the Clinical disorder themes *(see page 5)*

BSTA-2.03 Towards personalised medicine for antidepressant drugs: a machine learning approach

Primary Supervisor: Dr Daniel Stahl

Academic Department: Biostatistics and Health Informatics

Email: daniel.r.stahl@kcl.ac.uk

Website: https://kclpure.kcl.ac.uk/portal/daniel.r.stahl.html

Second Supervisor: Dr Raquel Iniesta

Academic Department: Biostatistics and Health Informatics

Email: raquel.iniesta@kcl.ac.uk

Website: https://kclpure.kcl.ac.uk/portal/raquel.iniesta.html

Project Description

Background: Individuals with depression differ in their response to treatment with antidepressants. It has been estimated that at least 40% of variation is attributable to genetic factors. Clinical factors showed a moderate role in predicting treatment response. The development of algorithms integrating genetic and clinical factors to predict treatment outcomes in depression may enable clinicians to select optimal medication for each patient.

Novelty and Importance: Developing algorithms for treatment personalization from large datasets is a challenging task for which traditional statistics had limited success. Problems arise when the number of predictors exceeds the number of individuals, there are missing data, variables are highly correlated and effect sizes are small.
This PhD studentship will develop new algorithms to personalise antidepressant treatment based on Topological Data Analysis (TDA) a set of Multivariate machine learning techniques (MML) that have successfully been applied to precision medicine studies, but not yet in psychiatry. Algorithms will be applied to genetic and clinical data from 3899 subjects from two clinical trials (GENDEP, STAR*D) and three observational studies (MARS, MUENSTER, PRN-AMPS). This research may produce clinically relevant predictive models to guide clinicians in antidepressant treatment selection.

Primary aim(s): To develop new MML algorithms that combine genetic with clinical variables to improve prediction of antidepressant treatment response at the individual level.

Planned research methods and training provided:

This PhD studentship will include an individualised training plan in methodology required for personalised medicine (e.g. statistical genetics, machine learning, longitudinal analysis of multivariate phenotypes and predictive modelling).

Objectives:

Year 1: Develop and implement free standing code of a method for predictors selection based on the mapper TDA algorithm.

Year 2: Develop and implement free standing code of a predictive algorithm based on the combination of a TDA method for predictors selection and a machine learning classifier.

Continued on next page

BSTA-2.03 Towards personalised medicine for antidepressant drugs: a machine learning approach

Year 3: Validate the predictive algorithm internally and externally, in GENDEP, STAR*D, MARS, MUENTSER and PRN-AMPS datasets. Apply the developed algorithms to assess the role that clinical and genetic variation plays in response to antidepressant treatment in GENDEP, STAR*D, MARS, MUENTSER and PRN-AMPS studies.

This project combines (a) the development of cutting-edge methods for treatment personalisation, using TDA as a novelty, plus (b) the practical application of the developed methods to a very relevant area, with a real potential to improve patients' health. We would welcome applications from students with a background in mathematics or other quantitative sciences (e.g. bioinformatics, computer science) and from biomedical with a strong commitment to developing statistical skills.

Two representative publications from supervisors:

1: Machine learning, statistical learning and the future of biological research in psychiatry. **Iniesta R, Stahl D**, McGuffin P. In: Psychol Med. 2016 Sep;46(12):2455-65. doi: 10.1017/S0033291716001367. Epub 2016 Jul 13.

2: Combining clinical variables to optimize prediction of antidepressant treatment outcomes. **Iniesta R**, Malki K, Maier W, Rietschel 4, Mors 5, Hauser 6, Henigsberg 7, Dernovsek M8, Souery D, **Stahl D**, Dobson R, Aitchison KJ, Farmer A, Lewis CM, McGuffin P, Uher R. J
In: Psychiatr Res. 2016 Jul;78:94-102. doi: 10.1016/j.jpsychires.2016.03.016. Epub 2016 Apr 1.

Keywords: Antidepressant; Personalisation; Machine learning; Topological Data Analysis; Genetics;

BRC Theme/s:
Affective Disorders and Interface with Medicine
Bioinformatics and Statistics
Biomarkers and Genomics