Language-grounded causal inference from electronic mental health records
PhD Studentship

We are pleased to invite applications for a doctoral degree in the Department of Psychiatry, in association with St Cross College of the University of Oxford. The 3-year doctoral degree will start in October 2026.

Estimating causal effects of psychiatric treatments from observational data is undermined by confounding variables that are rarely captured in structured records, but are richly documented in clinical free-text. This project addresses this challenge through three interconnected studies. First, it will distil knowledge from frontier large language models (LLMs) into small local language models (SLMs) deployable within secure NHS environments, using them to extract clinically relevant variables and phenotypes from mental health records. Second, it will compare SLM-extracted variables against expert judgement to determine which variables should enter causal estimation as confounders, and evaluate competing causal discovery algorithms across multiple datasets. Third, it will task SLMs with designing an emulated trial, benchmarking the result against an existing gold-standard emulation and against independent causal inference experts. Together, these projects will establish whether privacy-preserving language models can reliably support causal inference in psychiatry at scale.

The PhD (DPhil) programme will be supervised by Dr Andrey Kormilitzin (University of Oxford) and Professor Erin Evelyn Gabriel (University of Copenhagen) with support from the SMARTBiomed consortium. The DPhil will provide the student with multidisciplinary skills, training in language models, natural language processing, statistical machine learning and causal inference.

The scholarship will fund course fees up to the value of home fees*, a tax-free stipend of no less than £20,780 per annum), plus additional support for research expenses, conference attendance, and consumables.

*Students with overseas fee status should be advised that they would need to fund the remainder of their fees from alternative sources. In exceptional cases, a request for additional support may be made at the funder's discretion.

Language-grounded causal inference from electronic mental health records
PhD studentship

Applicants should have, or expect to gain, at least an upper second-class honours degree in a field relevant to medical science and with strong foundations in statistics or machine learning and hands-on Python experience with superior programming skills; prior knowledge of causal inference and clinical NLP are an advantage. The student will be jointly supervised across Oxford and Copenhagen, with access to large-scale psychiatric datasets and structured training in causal inference, clinical NLP, and secure AI deployment through the SMARTbiomed researcher development programme.

You will need to apply for this studentship via the main University online graduate application form, and pay an application fee of £20. The application form, all supporting materials required for the programme (including references) and payment must be submitted by the appropriate studentship deadline. To access the application form and application guide please visit our website at www.graduate.ox.ac.uk/apply.

Informal enquiries are encouraged, please email: andrey.kormilitzin@psych.ox.ac.uk
Deadline for submission of applications: 12.00 noon (UK time), Friday 10th April 2026
Interview date: w/c 25th May 2026
Studentship code for application: 26PSYCH06WEB