Title of PhD project	**Statistical and machine learning methods for prediction and causality in the presence of non-proportional hazards, with application to large databases of transplant patients**	
Supervisor	Professor Linda Sharples	LSHTM
Co-Supervisor	Professor Ruth Keogh	LSHTM
Co-Supervisor	Dr Kate Walker	LSHTM
Brief description of project	We are seeking an outstanding candidate for a PhD project that lies on the intersection between statistics, health data science and clinical medicine. Using large databases of patients undergoing transplantation the successful candidate will develop methods for dealing with time to event outcomes (e.g. death or organ failure), in the presence of one or more effects that vary over time (non-proportional hazards). Key aims are: • to provide clarity on the causal interpretation of multiple time-varying effects on outcomes, • to apply appropriate methods to prediction of medium to long-term outcomes with application to patients undergoing liver transplantation, using data from UK and North American transplant databases, • to assess important causal questions (in collaboration with clinical colleagues), for example, how much do loco-regional therapies affect survival of patients on the waiting list for transplantation. The ideal candidate will have an MSc or equivalent experience in statistics, health data science or a related field, although a four-year scholarship is also an option (including MSc training in the first year). We particularly welcome applications from candidates in underrepresented groups.	
Skills we expect a student to develop/acquire whilst pursuing this project	Expertise in the design and analysis of observational studies with time to event outcomes involving non-proportional hazards and/or competing risks.	

	Expertise in both statistical and machine learning methods for investigating causal mediation for time to event outcomes with non-proportional hazards and/or competing risks. Experience of working in a multi-disciplinary environment (clinical, statistical, data science) to produce relevant, usable and generalisable methodology in this context. Experience of translating complicated technical analysis to clinical, patient groups and other non-statistical colleagues.
Particular prior educational requirements for a student undertaking this project	An MSc in either Medical Statistics, Statistics or Health Data Science, or a related MSc with a substantial quantitative component (e.g. Applied Mathematics, Engineering) would be ideal. For UK students a first class or upper second class undergraduate degree in a subject with a substantial quantitative component (e.g. Maths, Statistics) is required. For non-UK students comparable academic training and/or experience is required.
Project key words	Statistics Machine learning Survival Causal mediation
Possible under 1+4 route? Master's options identified.	Yes LSHTM - MSc Medical Statistics LSHTM - MSc Health Data Science
MRC Core Skills developed through this project	Quantitative skills
MRC LID themes	Health Data Science
Further reading	Dynamic survival prediction combining landmarking with a machine learning ensemble: Methodology and empirical comparison Assessing the Impact of Suboptimal Donor Characteristics on Mortality After Liver Transplantation