Title of PhD project	**Developing methods to combine analytic results from small samples**	
Supervisor	Professor Joan Morris	SGUL
Co-Supervisor	James Carpenter	LSHTM
Co-Supervisor	Joachim Tan	SGUL
Brief description of project	There is a large increase in the availability of routinely collected electronic health care data from different health care settings across Europe. Many countries are unable to share data on individual cases, but may provide aggregate data and analytic results. Two studies in particular are using such analytic results: EUROlinkCAT is evaluating the health and educational achievements of children with specific congenital anomalies compared to children without the congenital anomalies across Europe and ConcePTION is evaluating the safety of medications in pregnancy. However, there is a lack of statistical methods to combine analytic results from different sources when the samples within each data source may be small, which occurs when analysing rare congenital anomalies and/or specific medications in pregnancy. Methods are being developed to combine data from Kaplan Meier survival curves (to analyse the survival of children with congenital anomalies) and to combine data from median values (to analyse the length of hospital stays). However these methods need to be refined further and additional methods concerning combining data with categorical outcomes would be useful.

The Observational Health Data Sciences and Informatics (OHDSI) program is a multi-stakeholder, interdisciplinary collaborative whose aim is to bring out the value of health data through large-scale analytics. All their solutions are open-source. They currently have an "EvidenceSynthesis" package (in R), which only includes the fitting of random and fixed effects meta-analysis.

The aim of this PhD is to continue the collaboration with researchers in EUROlinkCAT and ConcePTION to develop the methods they require and to also contact the OHDSI program to suggest updates for their "EvidenceSynthesis" | |

	package to include more tools for synthesising data, particularly when it is sparse.
Skills we expect a student to develop/acquire whilst pursuing this project	Ability to analyse data sets in R and to write open-source R scripts. Detailed knowledge about meta-analytic techniques. Knowledge of health care data sources across Europe as well as general knowledge about congenital anomalies and safety of medications in pregnancy will be acquired.
Particular <u>prior</u> educational requirements for a student undertaking this project	BSc and preferably MSc in mathematics/statistics/medical statistics
Project key words	Meta-analysis R programming Health Care Data
Possible under 1+4 route? Master's options identified.	Yes LSHTM – MSc Medical Statistics
MRC Core Skills developed through this project	Quantitative skills
MRC LID themes	Health Data Science
Further reading	[Evaluation of stability of directly standardized rates for sparse data using simulation methods](#) [Missing data: A statistical framework for practice](#)