

Detecting outbreaks of infections using highly dimensional data

Supervisors: Simon Spencer and Noel McCarthy

Outbreaks of infection have typically been identified when there is a sudden increase in the numbers of infections at a particular time, and in a particular place. Typically each infection in an outbreak shares a common source, even though some may have different transmission routes such as early infections coming from the environment followed by person to person spread. Sequencing of the genomes of microbes now occurs on a large scale. This allows a new dimension, genetic space, in which infections can be found to be related. Sometimes this allows identification of outbreaks that would otherwise have been missed, with a persistent source causing a large outbreak but spread over a long time and area so that it is missed by traditional methods.

Although there is extensive work on algorithms to detect temporal variation and exceedances there is far less on combining time, space and genomic dimensions of data.

A current MathSys student is currently completing a PhD developing Bayesian methods to jointly analyse these data with us. There are opportunities to extend this work further and to undertake application with a wide range of disease data from national infectious disease surveillance in England.

This will be in collaboration with Public Health England with access to national data and would offer opportunities for time working at Public Health England NIHR Health Protection Research Unit in Gastrointestinal Infections and the NIHR Health Protection Research Unit in Genomics and Enabling Data. Students would therefore have access to the NIHR academy and professional training offered by the NIHR (National Institute for Health Research).