PGR MINI-SYMPOSIUM

Advanced Quantitative Methods

MANCHESTER 1824
The University of Manchester

Lancaster University

UNIVERSITY OF LIVERPOOL

North West Doctoral Training Centre

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

The University of Manchester

3rd November 2015 10.30-17.00

Getting to the University of Manchester Campus

Travel By Car:

The University campus is located on Oxford Road, near to Manchester city centre. There is a public car park on the campus located to the rear of the Humanities Bridgeford Street building. Access to this carpark is from Booth Street West – please note that there is no access to the car park via Higher Cambridge Street. The postcode for the car park is M15 6PB. Payment for parking may be made by cash, debit card or credit card.

Directions to campus:

From M62 (eastbound), M602

Leave the M62 at J12 and join the M602.
At the end of M602, join Regent Road (A57) and continue along and join the A57M (Mancunian Way).
Leave at the second exit signposted A34 Congleton (hair-pin bend) and keep left, following signs for The University of Manchester.
Turn left onto the A34 (dual carriageway) and get in the right-hand lane.
Turn right at the first set of traffic lights into Grosvenor Street.
Stay in the left-hand lane and turn left at the next set of traffic lights onto Oxford Road (B5117).

From M62 (westbound), M60

Leave M62 at J18 and join M60 ring road.
Leave M60 at J22 and turn right at traffic lights onto Oldham Road (A62).
Continue along until the end and turn left to join Great Ancoats Street (A665).
Follow signs for The University of Manchester and join the Mancunian Way (A57M).
Take the exit from the Mancunian Way signposted A5103/Manchester Airport (M56)/Birmingham (M6).
At the roundabout take the first exit onto Higher Cambridge Street.
Turn left at the second set of traffic lights onto Booth Street West and then turn right onto Oxford Road (B5117).

From M6/M56

Leave the M6 at Junction 19 and turn right onto the A556.
Follow signs for Manchester Airport and join the M56.
Stay on M56 until it joins Princess Park Way (A5103).
Continue for a further four miles (6km) before turning right at Young's Royal Brewery into Moss Lane East (B5219).
Continue to the T-junction and turn left onto Oxford Road (B5117).
Proceed down Oxford Road and the University campus is straight ahead.

From M67

At the end of the motorway join Hyde Road (A57).
Continue along for approximately three miles (4.5 km), following signs for the city centre.
Upon reaching a major roundabout (the Apollo Theatre is on the left) take the second exit, Brunswick Street (signposted The University of Manchester).
Continue along and go straight through the next set of traffic lights.

Getting to the University of Manchester Campus

Travel by train

The University of Manchester is located close to both Piccadilly main line station and Oxford Road train station, with Victoria train station a little further away.

For details of timetables, tickets and other rail information, please ring National Rail Enquiries on 08457 48 49 50 (+44(0)20 7278 5240 from overseas) or visit:

National Rail Enquiries website

Please see the links below for a map with directions from Victoria Station and Piccadilly Station.

Map and directions from Victoria Station to the University Campus

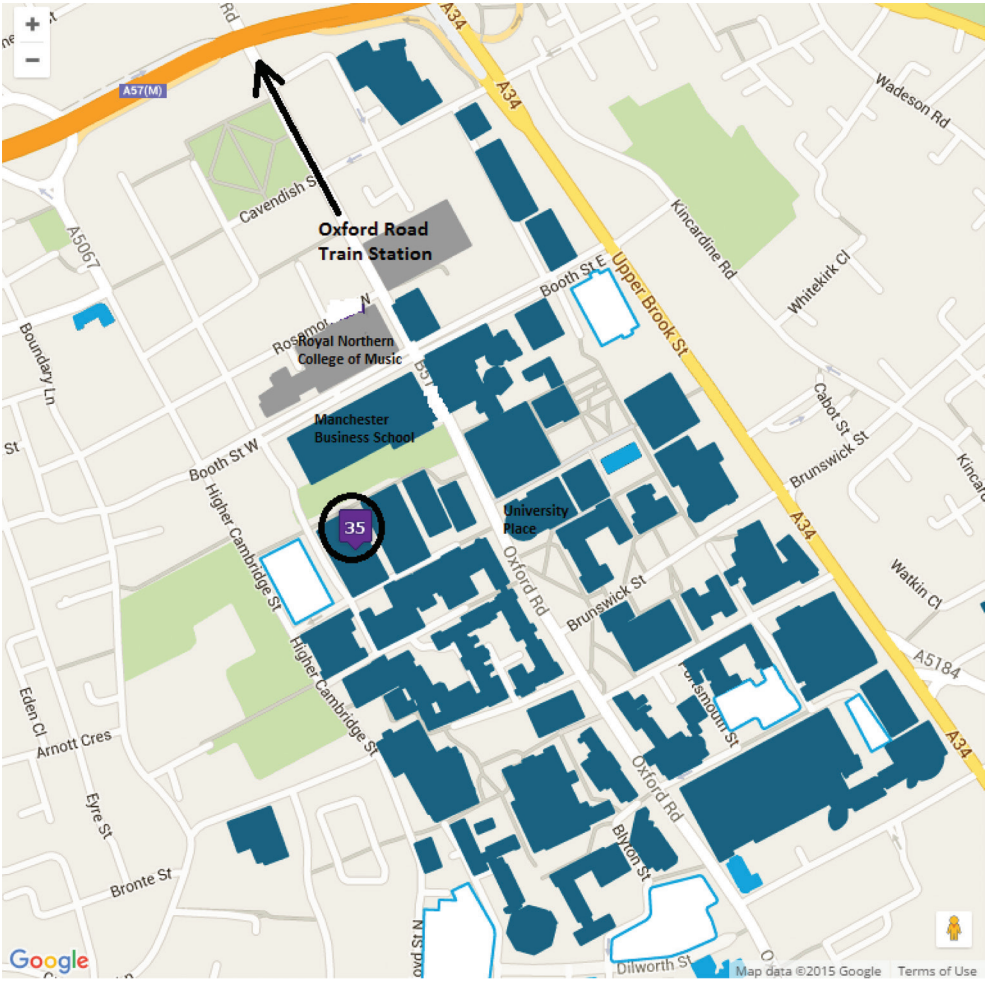
Map and directions from Piccadilly to the University Campus

A taxi from Piccadilly Station to the Oxford Road area of the campus will cost you about £5.
Oxford Road station to Oxford Road campus

To get to the Oxford Road area of the campus, go down the station approach to Oxford Road and turn right. The campus is a ten to fifteen minute walk. Alternatively, buses stop outside the Palace Hotel.

Symposium Location

The PGR Mini-Symposium will be held in Lecture Theatre G33 of the Humanities Bridgeford Street Building on the Oxford Road Campus of the University of Manchester. The Humanities Bridgeford Street Building is numbered 35 on the map beneath.



Timetable

AQM Mini-Symposium Schedule 2015

10.00-10.30	Arrival and Registration
10.30-10.35	**Welcome**
Session 1	*Chair: Sebastian Franke*
10.35-10.55	Deserving and Undeservingness Claimants in Diverse Welfare States: Results from Survey Experiments Administered in Britain and the Netherlands *Anouk Kootstra*
10.55-11.15	Measuring models for understanding the social challenges of caring for the elderly in Brazil and the UK *Rebecca Pattinson*
11.15-11.45	**Coffee**
Session 2	*Chair: Michael Chipeta*
11.45-12.05	Putting "Geo" into Geodemographics: Evaluating the Importance of Spatial Proximity and Context for National Classification Performance *Alexandros Alexiou*
12.05-12.25	Partnership status, health and mortality: protection or selection? *Sebastian Franke*
12.25-12.45	An Agent-Based Model approach to understanding changes in ethnic relations: applications to neighbourhood ethnic composition and health *Frensis Bras*
12.45-14.00	**Lunch**
Session 3	*Chair: Rebecca Pattinson*
14.00-14.20	Geography and Social Distribution of Malaria in Indonesian Papua *Wulung Hanandita*
14.20-14.40	Modelling spatial variation in survival among HIV patients diagnosed with KS in Zomba, Mala-wi *Emmanuel Singogo*
14.40-15.00	Adaptive Geostatistical Design and Analysis for Sequential Prevalence Surveys. *Michael Chipeta*
15.00-15.30	**Coffee**
15.30-15.45	Open discussion
15.45-16.35	**Keynote:** From PISA to captain Picard: the impact of measurement error in the social sciences *Alexandru Cernat*

Keynote Speaker

Alexandru Cernat

Alex studied for his PhD in survey methodology at the University of Essex, Institute for Social and Economic Research on the topic of mixed modes in longitudinal data. His main research interest is on the use of latent variable modelling to estimate measurement error with a special focus on survey methodology and longitudinal designs. He is currently a Research Associate at the Cathie Marsh Institute for Social Research and the National Centre for Research Methods where is working on dealing with missing data in biomarkers.

From PISA to captain Picard: the impact of measurement error in the social sciences

Measurement error is a pervasive issue in the social sciences, affecting most of the quantitative data used: surveys, admin data and even "big data". While most researchers are aware of the issue it is nevertheless usually ignored. The presentation will discuss the main types of measurement error found in survey data, how they can impact substantive results and what can applied researchers do about it. It will conclude with two examples of high "impact" survey research and how their results can be misleading when measurement error is ignored.

Session 1 - *Chair: Sebastian Franke*

Deserving and Undeservingness Claimants in Diverse Welfare States: Results from Survey Experiments Administered in Britain and the Netherlands

Anouk Kootstra, The University of Manchester *anouk.kootstra@manchester.ac.uk*

Many scholars have observed how increasing immigration may threaten the legitimacy of west-European welfare states (Alesina and Glaeser 2004; Kymlicka and Banting 2011). Although many studies examine this relationship on a macro-level, research that investigates the effects of migration and ethnic diversity on public welfare attitudes on the micro-level is scarce (Ford 2015). Is public generosity towards welfare claimants, such as pensioners, sick and disabled people and the unemployed, dependent upon the migration history or ethnic background of those claimants? And why, then, are migrant and ethnic minority claimants considered less deserving than majority claimants?

Previous research in the field suffers from several theoretical and methodological shortcomings that might have resulted in over- and/or underestimation of the prevalence of welfare chauvinism and welfare ethnocentrism. I present new measures of the phenomena that are based upon an innovative experimental design that overcomes the drawbacks of previous measures. In my presentation I will show the results of an experiment administered across five waves of newly gathered panel data among 5000 respondents in Britain and the Netherlands. This setup combines the experimental, longitudinal and comparative research designs in an unprecedented way, and makes both theoretical and methodological contributions to the field.

My presentation reflects these different objectives: the first part introduces these new measures of welfare chauvinism and ethnocentrism, discusses their validity and compares them to traditional measures. The second part makes a substantial contribution by demonstrating the prevalence of welfare chauvinism and ethnocentrism in Britain and the Netherlands, and by showing that public resistance against migrant or ethnic minority welfare claimants can be explained by a disparity in the deservingness rhetoric: whereas majority claimants are generally perceived as 'unlucky' and hence deserving of financial support, migrants and ethnic are more often seen as 'lazy' and therefore undeserving of receiving welfare benefits.

References

Alesina, Alberto, and Edward L. Glaeser. 2004. Fighting Poverty in the US and Europe. Oxford: Oxford University Press.

Ford, Robert. 2015. "Who Should We Help? An Experimental Test of Discrimination in the British Welfare State." Advanced Access, Political Studies.

Kymlicka, Will, and Keith Banting. 2011. "Immigration, Multiculturalism, and the Welfare State." Ethics & International Affairs 20 (03): 281–304.

Measuring models for understanding the social challenges of caring for the elderly in Brazil and the UK

Becky Pattinson, Lancaster University *r.pattinson2@lancaster.ac.uk*

The increasing demand for care for the elderly in an ageing population is a problem faced by many countries, including England and Brazil. We aim to compare the care of the elderly in England and Brazil using the three domains of health, socialisation and well-being, as per the definition of active ageing by the World Health Organisation. To analyse the elderly in Brazil we used data from the 2008 PNAD (National Household Sample Survey) which used a complex sampling plan and contained a hierarchical structure that required consideration. The potential levels in the hierarchical structure were individuals, household, sector, municipality and strata. We used multilevel modelling and analysed the VPC of the data to determine which of the levels in the hierarchical structure to include in our analysis. As part of our analysis of the health of the elderly in Brazil, multilevel ordinal logistic regression was used to model their self-rated health to determine the effect of various demographic and socioeconomic factors, whilst accounting for the hierarchical structure in the data.

Session 2 - *Chair: Michael Chipeta*

Putting "Geo" into Geodemographics: Evaluating the Importance of Spatial Proximity and Context for National Classification Performance

Alexandros Alexiou, University of Liverpool *A.Alexiou@liverpool.ac.uk*

A Geodemographic analysis is a methodology that classifies patterns of socio-economic and built environment attributes of small area geography. It involves the process of creating classifications that represent categorical summary measures of neighbourhoods on the basis of the characteristics of its residents. The inferential nature of the aggregations relies on the notion of societal homophily, commonly addressed as the "birds of a feather flock together" phenomenon.
A particular issue within geodemographics is that classifications lack any explicit specification of geographic context within the clustering process. Current techniques smooth away differences across geographic zones which could obscure interesting local patterns. This research sets out to explore the issue of geographic sensitivity, and particularly how localised classifications deviate from national socio-spatial patterns, and how this differentiation can be measured effectively. For this purpose, a set of fixed input attributes for Output Area zonal geography is used to build classifications with different geographic extents. A number of scales are considered (local, regional, national) as attribute contextual weights to demonstrate the impact on final classification outcome when input variables are kept constant. An initial exploration is presented through the case studies of Liverpool, Manchester and Leeds, and based on the preliminary findings a methodology is introduced to calculate the variation of localised socio-spatial patterns to national standards.

Outcomes illustrate the degree of spatial autocorrelation across neighbourhoods in terms of socio-spatial pattern dissimilarity, prompting the existence of distinctive drivers and constraints on the mechanisms of this variation. For simplicity, administrative definitions of context have been used for this study, however, it is recognised that these may not represent true functional localities, and as such, further work is required about how local extents might be defined, and what impact this geography will have on final classifications.

Partnership status, health and mortality: protection or selection?

Sebastian Franke, University of Liverpool *S.Franke@liverpool.ac.uk*

Research on health and mortality by marital status shows lower mortality rates and better health for married persons in comparison to single and separated individuals. Those differences, usually stronger for men than for women, persist even when controlling for socio-demographic and economic characteristics of individuals. Recent changes in England and Wales over the last 40 years - such as the rise in cohabitation, divorce rates, lone parent families, and life expectancy; as well as a general change in household structures – invite a re-evaluation of these differences by focusing on health and mortality by different living arrangements. The aim of the project is to analyse the trends in mortality differences by partnership status in England and Wales. It investigates the effect of partnership status on health as well as the impact of health on partnership status. Further, it will demonstrate if and how much the increase in cohabitation leads to a decrease in th e mortality difference between cohabitants, cohabiting couples and married couples. By applying hazard models to data from the ONS Longitudinal Study (ONS LS) between the 2001 and the 2011 censuses, the project investigates health and mortality by partnership and family status

and examines the causes of mortality differentials. A more in-depth analysis will be undertaken using the British Household Panel Study data (BHPS). Multi-level equation survival models will be used to control for selection in and out of partnership statuses.

An Agent-Based Model approach to understanding changes in ethnic relations: applications to neighbourhood ethnic composition and health

Frensis Bras, University of Manchester frensis.bras@manchester.ac.uk

Ethnic minorities who live in areas with a higher concentration of ethnic minority residents have better health outcomes compared to their peers living in relatively white areas. This is called the ethnic density effect. One of the ways in which ethnic density might contribute to better health outcomes is by reducing the amount, and effect of racism. This might be because areas that are diverse harbour norms and values that are not accepting of racist behaviour, or because they contain fewer perpetrators. Due to a lack of available data and the difficulty of accounting for confounding and endogenous effects in observational studies, the exact mechanisms driving the association between ethnic density and the reduction in racism have not been adequately tested. I aim to shed light on the aforementioned effects by means of an Agent-Based Model (hereafter ABM). ABM is a computational method that generates patterns of group behaviour on the basis of individual behav iour. This relatively novel technique can take into account the complex ways in which communities develop over time and help to untangle the different causal relationships. ABM might be particularly applicable to researching the relationship between ethnic density and racism because of the diverse and dynamic effects that are at play, the lack of suitable data, and the way in which ABM can incorporate spatial and temporal elements. Findings up to date indicate that increased ethnic density does create an environment with less prejudice and fewer perpetrators, however this relationship might be influenced by other factors such as segregation and deprivation.

Session 3 - *Chair: Michael Chipeta*

Geography and Social Distribution of Malaria in Indonesian Papua

Wulung Hanandita, The University of Manchester wulung.hanandita@postgrad.manchester.ac.uk

Despite being one of the world's most endemic areas, only little is known about the social and spatial distributions of malaria in Indonesian Papua. Existing studies tend to be descriptive in nature; their inferences are prone to confounding and selection biases. At the same time, there remains limited malaria-cartographic activity in the region. Analysing a subset (N = 22,740) of the National Basic Health Research 2007 dataset (N = 987,205), this paper aims (1) to quantify the district-specific risk of malaria infection in Papua and (2) to understand how socio-demographic/economic factors measured at individual and district levels are associated with one's probability of contracting the disease. Methods: We adopt a Bayesian hierarchical logistic regression model that accommodates not only the nesting of individuals within the island's 27 administrative units but also the spatial autocorrelation among these locations. Both individual and contextual characteristics are included as predictors in the model; a normal conditional autoregressive (CAR) prior and an exchangeable one are assigned to the random effects. Robustness is then assessed through sensitivity analyses using alternative hyperpriors. Results: We find that, net of confounding and spatial effects, rural Papuans as well as those who live in poor, densely forested, lowland districts are at a higher risk of infection than their counterparts. We also find age and gender differentials in malaria prevalence, if only to a small degree. Finally, we present a model-based risk map that is invaluable for guiding efficient and equitable allocation of resources available for disease control. Conclusions: As much as being geography-dependent, malaria in Indonesian Papua also appears to be a disease of poverty. This means that malaria eradication requires not only biological (proximal) interventions but also social (distal) ones. On top of that, the present study also demonstrates the utility and feasibility of Bayesian methods in supporting malaria eradication efforts. Keywords: malaria, map, papua, indonesia, bayesian, spatial, multilevel

Modelling spatial variation in survival among HIV patients diagnosed with KS in Zomba, Malawi.

Emmanuel Singogo, Lancaster University e.singogo@lancaster.ac.uk

Use of spatial methods to study geographical variation in survival patterns can help to reveal underlying confounding factors and in identifying regions requiring more at- tention. In epidemiology and biostatistics, there has been a growing interest for the last two decades in developing different analysis methods to use spatial methods for survival data. These methods have been used extensively in cancer and HIV epidemiology. In our study, we employ the use of a Weibull spatial frailty model to study variation in survival among HIV patients with Kaposi's sarcoma (KS). Our objectives are to understand the epidemiology of KS among HIV patients and whether the survival outcomes are char- acterized by the geographical distribution. The initial exploratory analysis indicates presence of significant heterogeneity among patients. The results from the spatial model show that up to 47% of possible residual spatial variation in survival after accounting for patient-specific covariates; age, sex, year of registration, TB status, distance to the HIV clinic and occupation. We also present a probabilistic map showing areas with high and low risk of dying (HR $\geq$ 1.5). These differences in hazards highlight the underlying heterogeneity in the population and could be further investigated to inform targeted treatment and care. Adaptive Geostatistical Design and Analysis for

Sequential Prevalence Surveys

Michael G. Chipeta, Lancaster University m.chipeta@lancaster.ac.uk

Introduction

Geostatistical methods are used increasingly to support disease control efforts and analyse heterogeneity in disease prevalence. Adaptive geostatistical sampling designs (AGD) allow collection of exposure and outcome data over time to depend on information obtained from previously collected data so as to optimise data collection towards the analysis objective. AGDs are especially useful in resource poor settings where uniformly precise mapping may be unrealistically costly and the priority is often to identify critical areas where interventions can have the most health impact. If successfully implemented, AGDs should outperform current gold standard sampling by improving on predictive performance and hotspots identification.

Methods

We previously assessed the efficiency of specific batch AGDs relative to their singleton adaptive and non-adaptive counterparts by using simulations. In most practical settings however, batch sampling is more realistic than singleton sampling. In batch sampling, locations are chosen in clusters of size $b>1$, allowing a new batch, $\{x(kb+1),\ldots,x(k+1)b\}$, to depend on data obtained at locations $x1,\ldots,xkb$. We introduced a batch adaptive sampling approach to an on-going rolling household level Malaria Indicator Survey (rMIS) that is being conducted within a large scale, five-year malaria transmission reduction project (The Majete Malaria Project) in communities living around the Majete Wildlife Reserve, Malawi.

Results

We will present how this real-life application performs to describe local variation in malaria infection in under-five children, identify hotspots that could guide more targeted disease control efforts and investigate the association of prevalence with environmental. We will present this using a combination of data collected during the survey and publicly available, remotely sensed climate and environmental information.

Conclusion

With the increasing use of electronic data capture that allows real-time analysis of incoming data, and advances in spatial statistical techniques, novel adaptive sampling methods could become instrumental in monitoring and accelerating disease transmission reduction.

APPLICATION: EBOLA EPIDEMIC, LIBERIA 2014

- Line-listing data from the large Ebola Virus Disease (EVD) outbreak in Liberia 2014 was analysed
- Epidemiological information included date of onset of symptoms and the residential district of each case
- There was significant heterogeneity in outbreak timings and epidemic size between districts (Figure 2, 3)
- Districts far apart were likely to have different transmission patterns due to factors such as different human movement patterns and environmental conditions
- We applied our spatial method to this data, using information on the serial interval (the time taken for each case to infect further cases) and assumed human mobility patterns, based on a simple gravity model (Figure 4)

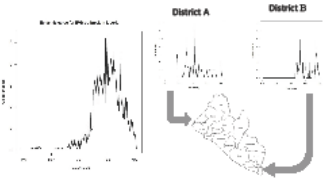


Figure 3. Epidemic curve for whole of Liberia, and epidemic curves for select districts. At the district level, the pattern is different indicating an underlying heterogeneity in transmission dynamics.

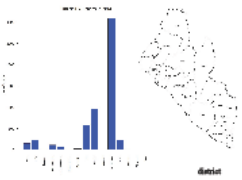


Figure 2: Distribution of Ebola cases in Liberia by district, and map of district locations

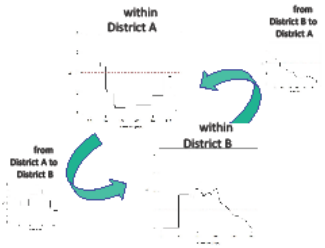


Figure 4. estimates for within and between districts transmission of infection

SUMMARY

- Extension of reproduction number estimation to include spatial information.
- New tool for public health – real-time assessment of interventions.

FUTURE WORK

- Developing method to use the partial likelihood approach to incorporate covariates in the estimation of R_{it}, permitting the impact strength of interventions and other covariates to be estimated

ACKNOWLEDGEMENTS

Incorporating spatial aspects into reproduction number estimation for infectious disease outbreaks

James Chirombo[1,2], Peter J. Diggle[1], Anja Terlouw[2], Jonathan M. Read[1]

1. CHICAS, Lancaster Medical School, Lancaster University, UK
2. Malawi Liverpool Wellcome Trust Clinical Research Programme, Blantyre, Malawi

BACKGROUND

- Controlling outbreaks of infectious disease is often achieved through non-pharmaceutical public health interventions, such as quarantining, contact tracing, or curfews
- Quantitative tools are needed to assess the growth of epidemics and efficacy of interventions
- Effective reproductive number (R_{it}) is the average number of secondary cases caused by a each case at a calendar time period, during and epidemic
- Estimating R_{it} during an outbreak is important for assessing the progress of an epidemic and for gauging the impact of interventions
- If $R_{it} < 1$, the epidemic is under control
- If $R_{it} > 1$, the epidemic is still growing and out of control
- Public health interventions aim to bring R_{it} below 1

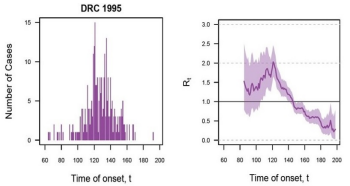


Figure 1. An example of R_t estimated for a small Ebola virus outbreak in Democratic Republic of Congo, 1995. The epidemic is under control after 140 days since outbreak onset.

MOTIVATION

- There are several limitations of the current methods used to estimate Rt. One major limitation is a failure to incorporate available spatial information: R_{it} is assumed to vary with time only. Current methods calculate a global R_{it} estimate for a spatial region, ignoring localized effects.
- Spatial heterogeneities in transmission are very common for outbreaks in many settings (Figure 2).
- By incorporating spatial information, transmission rates at specific locations may be estimated, and 'hot-spots' sources of infection may be identified, therefore improving targeting control for scenarios where there is limited control resources.

AIMS

- To develop a method to estimate R_{it} which incorporates spatial information on cases
- Spatial information may be geographic location of a case, contact information (such as from contact tracing studies) describing how cases are related, or co-location information (such as which cases shared a ward, for hospital-based outbreaks)