GEOMED
2019 GLASGOW

Book of Abstracts

XI INTERNATIONAL, INTERDISCIPLINARY CONFERENCE ON SPATIAL STATISTICS, GEOGRAPHICAL EPIDEMIOLOGY AND GEOGRAPHICAL ASPECTS OF PUBLIC HEALTH.

University of Glasgow | School of Mathematics & Statistics

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

JOHNS HOPKINS
BLOOMBERG SCHOOL of PUBLIC HEALTH

CRC Press
Taylor & Francis Group
CHAPMAN & HALL

TUNNOCK'S CARAMEL WAFER

Contents

1. Abstracts for plenary talks and workshops

Workshop 1 - R and GIS, or R as GIS: handling and analysing spatial data

Roger Bivand
Norwegian School of Economics

Workshop goal
Provide enough understanding of how spatial (and maybe spatio-temporal) data is (and will be) represented in R.
First part
The representation of spatial and spatio-temporal data is important to its handling and preparation for analysis. While the sp package has provided legacy representations still used by many, the sf package introduces more modern spatial vector representations, and the stars package is moving towards a similar modernisation of raster array representations. As yet no replacement for the spacetime package has been proposed, although stars does reach out towards spatio-temporal data. The first session will also describe ongoing changes in external sofware used for representing spatial reference systems.
Second part
GIS and R: bridges or R as GIS? The second session will present general principles of GIS in R - R as GIS is becoming more feasible as new contributed packages appear, and/or R bridges to GIS to permit GIS operations to be done in GIS but statistical operations in R. The seesion will conclude by presenting ongoing improvements in thematic maps in the tmap and cartography packages, and interactive mapping in mapview.

Workshop 2 - Spatial and Spatio-temporal Bayesian Models with R-INLA

Marta Blangiardo[1] and Michela Cameletti[2]
[1]Imperial College London
[2]University of Bergamo

Workshop goal
Provide an introduction to the Integrated Nested Laplace Approximation approach (INLA) for the analysis of spatial and spatio-temporal data at the small area level.
First part

We will go through the basic of INLA for Bayesian inference and will then see how to model hierarchical structures, in particular focusing on spatially structured random effects through conditional autoregressive specifications; we will then extend the approach to include temporal dependency and touch briefly on spatio-temporal interactions.

Second part

After the coffee break participants will have the opportunity of running a tutorial on their computers, implementing some of the models seen in the first part of the workshop. It is recommended that people attending are familiar with R (https://www.r-project.org/) and with the basic of the Bayesian approach. It would be beneficial if attendees could bring their laptop with the latest version of R and INLA installed.

Places and health: history, concepts, and emerging directions

`Ana V. Diez Roux`
`Drexel University`

This plenary will provide an overview of epidemiologic work on places and health over the past 25 years, identifying key conceptualizations, summarizing main methodologic approaches and their limitations, and identifying promising areas for future exploration. It will reflect on what we have learned, what questions remain unanswered, and what emerging areas present new opportunities for the field.

Design and Analysis of Elimination Surveys for Neglected Tropical Diseases

`Peter Diggle`
`Lancaster University`

Neglected tropical diseases are “a diverse group of communicable diseases that prevail in tropical and subtropical conditions in 149 countries... Populations living in poverty, without adequate sanitation and in close contact with infectious vectors and domestic animals and livestock are those worst affected” (WHO web-site). International efforts to control these diseases have led to considerable reductions in prevalence to the point where, for many NTDs, their elimination as a public health problem has become a realistic proposition.

A strategy to confirm (or not) that elimination has been achieved requires consideration of (at least) the following issues.

1. What is the definition of elimination? Typically, this is that average prevalence over a defined geographical region, A, does not exceed a specified threshold, but other definitions are possible. For example, a low average prevalence could conceal isolated hot-spots of residual infection.
2. What is the appropriate choise for A? Typically, this will be an administrative region corresponding to a geographical unit on which a policy decision will be made (for example, cessation of mass treatment with prophylactic medication).

3. What level of uncertainty is acceptable to those responsible for making policy decisions? In low resource settings, where population-wide testing is impractical, some level of uncertainty is inevitable.

4. What data should be collected, and how should they be analysed?

In this talk, I will argue that the correct inferential framework for this problem is neither estimation nor testing, but rather prediction. This point of view leads to the following working definition of elimination, with obvious variations. Let $n(x)$ be the population density at location x and $p(x)$ the probability that an individual at x has the disease in question. The population-weighted average prevalence over A is

$$AP = \int_A p(x) dn(x),$$

and elimination is achieved if, given user-specified values for two positive constants, c and q,

$$\text{Prob}(AP < c) > q.$$

It follows that the goal of the study-design should be to deliver as tight as possible a predictive distribution for AP when its true value is close to c, subject to context-specific practical constraints. In my experience, this requires a statistical model for spatially referenced prevalence data that takes account of both measured covariate effects and unexplained stochastic spatial variation in $p(x)$.

`Acknowledgment`
Work-in-progress funded by the Gates Foundation through a grant to the Neglected Tropical Diseases Modelling Consortium led by Prof Deirdre Hollingsworth (University of Oxford).

Spatial precision, statistical precision – having the best of both worlds

`Robert Haining`
`University of Cambridge`

Developments in the collection, storage, retrieval and display of social, economic, health and environmental data enable us to observe and study our world in ever finer geographical detail. With increasing temporal frequency we are also better able to observe and measure changes taking place over short time spans. But to reap the benefits of spatial and spatial-temporal data precision we need to be able to obtain reliable estimates of attribute properties for these small spatial and spatial-temporal units. Unreliable estimation of variability associated with small area estimates undermines the benefits that can be derived from such data. The question is then: how can we have the best of both worlds – spatial and spatial-temporal precision and statistical precision?

This presentation will overview four main challenges in reaching this goal through data modelling: dependency, heterogeneity, data sparsity and uncertainty. We then discuss two modelling approaches - hierarchical modelling and spatial econometric modelling – and also consider the benefits of adopting a Bayesian (rather than frequentist) approach to inference. We conclude by looking briefly at some examples of both types of data modelling that illustrate some of the opportunities that arise in the areas of ecological modelling, infectious disease modelling and space-time disease forecasting.

2. Abstracts for talks

Spatial distribution of wasting, stunting, and underweight among children aged 5 and under in Kenya: A shared components modelling approach

Thomas Achia
University of the Witwatersrand

Background
Malnutrition is estimated to contribute to more than one third of all child deaths worldwide. This study is aimed at assessing the joint spatial variability in wasting, stunting and underweight in Kenya so as to provide a model-adjusted maps of malnutrition for Kenya.

Methods
Cross-sectional data from 19,005 children aged under-5 was obtained from the 2014 Kenya Demographic and Health Survey. Measurements of stunting, height-for-age z-score (HAZ) less than -2; wasting, weight-for-height z-score (WHZ) < -2 and underweight, weight-for-height z-score (WAZ) < -2 were computed. These data were aggregate to county level then the Moran's I statistic was used to assess the presence spatial autocorrelation. Bayesian spatial shared components models were fitted in WinBUGS, measuring and ranking areas according to specific and shared risk of malnutrition.

Results
The prevalence of underweight, stunting and wasting were 4.4%, 25.9% and 11.1%, respectively. Wasting (Moran's I=0.72, p<0.001) and underweight (Moran's I=0.56, p<0.001) exhibited significant spatial clustering and also had significant joint spatial clustering (Bivariate Moran's I=0.61, p<0.001). The results indicate that for both wasting [Posterior Risk (PR)=0.50, 95% Credible Interval (Cred. Int.)=0.495-0.515] and under-weight [PR=0.50, 95% Cred. Int.=0.491-0.511], about half of the total between-area variation in risk is captured by the shared component. The shared component has a slightly stronger association with risk of wasting than with risk of underweight (RR=1.59, 95% Cred. Inter=1.37-1.87).

Conclusions
Our results suggest a need for comprehensive programming for malnutrition interventions in the high risk Northern parts Kenya.

Using an agent-based model to investigate the impact of neighbourhood settings on children's physical activities

Jonatan Almagor
University of Glasgow

Emerging evidence in the health literature suggests that the built environment and the design of urban areas can facilitate healthier lifestyles and contribute to reducing the risk of noncommunicable disease. For children, the context of the neighbourhood settings, where they spent most of their time, may be especially important. A higher density of residential dwellings and a greater mix of local land uses have been found to increase the likelihood of children walking, cycling, and encourage local social interactions. Walkable neighbourhoods, with well-connected streets and low traffic volumes, are more likely to promote active travel to school. Land uses associated with nature appear to favour children's cognitive and physical development, and may positively affect emotional aspects, cognitive functioning and physical activity. The distribution and configuration of different land uses produces spatial heterogeneity in the built environment: in most cases, no one place resembles another. Accessibility and proximity to different land uses and facilities varies significantly depending on location; therefore, considering the implications of the spatial context is crucial when examining the impact of the neighbourhood setting on children's activities and physical activity. I present an agent-based model that simulates the daily activity of children during week days in various neighbourhoods across a city. The model examines the possible impact of the neighbourhood environment on activities children may engage, and the implications on levels of psychical activity.

A geostatistical framework for combining spatially referenced prevalence data from multiple diagnostics

Benjamin Amoah
Lancaster University

Multiple diagnostic tests are often used due to limited resources or because they provide complementary information on the epidemiology of a disease under investigation. Existing statistical methods to combine prevalence data from multiple diagnostics ignore the potential over-dispersion induced by the spatial correlations in the data. To address this issue, we develop a geostatistical framework that allows for joint modelling of data from multiple diagnostics by considering two main classes of inferential problems: (1) to predict prevalence for a gold-standard diagnostic using low-cost and potentially biased alternative tests; (2) to carry out joint prediction of prevalence from multiple tests. We apply the proposed framework to two case studies: mapping Loa loa prevalence in Central and West Africa, using miscroscopy and a questionnaire-based test called RAPLOA; mapping Plasmodium falciparum malaria prevalence in the highlands of Western Kenya using polymerase chain reaction and a rapid diagnostic test. We also develop a Monte

Carlo procedure based on the variogram in order to identify parsimonious geostatistical models that are compatible with the data. Our study highlights (i) the importance of accounting for diagnostic-specific residual spatial variation and (ii) the benefits accrued from joint geostatistical modelling so as to deliver more reliable and precise inferences on disease prevalence.

Spatial clustering of risks in disease mapping

Craig Anderson
University of Glasgow

Spatio-temporal disease mapping focuses on estimating the spatial pattern in disease risk across a set of non-overlapping areal units over a fixed period of time. The goal is to identify areas with a high average level of disease risk or where disease risk increases, in order to allow public health interventions to be focused on these areas. These aims are well suited to the statistical approach of clustering, and while much research has been done in this area in a purely spatial setting, only a handful of approaches have focused on spatio-temporal clustering of disease risk. We outline a novel model-based approach for clustering spatio-temporal disease risk data, by clustering areas based on both their mean risk levels and the behaviour of their temporal trends. The method is illustrated by a study of respiratory disease risk in Glasgow, Scotland.

Detecting Spatial Clusters of Disease Infection Risk Using Sparsely Sampled Social Media Mobility Patterns

Renato Assunção
Universidade Federal de Minas Gerais

Standard spatial cluster detection methods used in public health surveillance assign each disease case a single location (e.g., patient's home address), aggregate locations to small areas, and monitor the number of cases in each area over time. However, this approach lacks the accuracy and specificity to deal with infectious disease outbreaks where human mobility plays a key role. Here, we use social media data to capture the individuals mobility. We propose two new spatial scan methods (the unconditional and conditional spatial logistic models) which search for spatial clusters of increased infection risk in mobility patterns by maximizing a generalized log-likelihood ratio statistic over subsets of the data.

The methods correctly account for the multiple, varying number of spatial locations observed per individual, either by non-parametric estimation of the odds of being a case or by matching case and control individuals with similar numbers of observed locations. By applying our methods to synthetic and real-world scenarios, we demonstrate robust performance on detecting spatial clusters of infection risk from mobility data, outperforming competing baselines.

Evaluation of spatial and spatio-temporal variation in distribution and determinants of female genital cutting among girls aged 0-14 years in Nigeria, 2003-2017

Glory Atilola[1], N Kandala[1,2], G Li[1], C Nnanatu[1], P Komba[1], L Mavatikua[1], D Matanda[3], Z Moore[4], O Otibho[5], C Kabiru[3], E Lwanga[3], B Shell-Duncan[7], and G Mackie[6]
[1]Northumbria University
[2]University of the Witwatersrand
[3]Population Council, Kenya
[4]Independent Consultant, Population Council
[5]Population Council, Nigeria
[6]University of San Diego
[7]University of Washington

Reports from nationally representative surveys showed little progress towards the United Nations' FGM/C abandonment efforts in Nigeria, as prevalence in girls 0-14 years increased from 19% in 2003 to a peak of 30% in 2008 among eldest daughters, while a 0.9% increase was reported between 2013 and 2017 in all girls aged 0-14 years. We proposed a unified framework to understand geographic and temporal variations in the distribution and determinants of FGM/C risks among girls aged 0-14 years in Nigeria using the proximate determinantframework. In this framework,underlying determinants operate through behavioral determinants at individual, household and community levels. We utilized data from demographic and health surveys (2003-2013) and the multiple indicator cluster surveys (2007-2017). Flexible Bayesian geo-additive models accounted for the effect of time, individual-, household-and community-level factors. Results showed that the key determinants of FGM/C include religion, ethnicity, mother's FGM/C status(OR: 13.58; 95%CI:11.97-15.44)and support for continuation of FGM/C(OR:11.94; 95% CI:10.81, 13.29).Another finding is the increased likelihood of a girl's cutting with low education status and age of mother. We found no evidence of association between media exposure and risk of cutting in Nigerian girls. Similar pattern of geographic variation in risk was found between the two periods with increased risk in Northern Nigeria and a decline in the South. This may be due to low uptake of existing interventions in northern Nigeria mostly predominant in the southern part of the country. Future research will explore the interaction dynamics of the effects of geography and time in the evolution of FGM/C in Nigeria.

Development of the Australian Neighbourhood Social Fragmentation Index and its association with spatial variation in depression across communities

Nasser Bagheri
The Australian National University

Background

We know little about how community structures influence the risk of common mental illnesses. This study presents a new way to establish links between depression and social fragmentation, thereby identifying pathways to better target mental health services and prevention programs to the right people in the right place.

Method

A principal components analysis (PCA) was conducted to develop the proposed Australian Neighbourhood Social Fragmentation Index (ANSFI). General practice clinical data was used to identify cases of diagnosed depression. The association between ANSFI and depression was explored using multilevel logistic regression. Spatial hot spots (clusters) of depression prevalence and social fragmentation at the statistical area level 1 (SA1) were examined.

Results

Two components of social fragmentation emerged, reflecting fragmentation related to family structure and mobility. Individuals treated for depression in primary care were more likely to live in neighbourhoods with lower socioeconomic status and with higher social fragmentation related to family structure). A 1-SD increase in social fragmentation was associated with a 16% higher depression prevalence (95% CI: 11%, 20%). However, the association attenuated with adjustment for neighbourhood socio-economic status. Considerable spatial variation in social fragmentation and depression patterns across communities was observed.

Conclusions

Developing a social fragmentation index for the first time in Australia at a small area level generates a new line of knowledge on the impact of community structures on health risks. Findings may extend our understanding of the mechanisms that drive geographical variation in the incidence of common mental disorders and mental health care.

Outfoxing rabies: a case study for fitting non-separable space-time log-Gaussian Cox models using R-INLA

Laurie Baker
University of Glasgow

Many processes are strongly linked in space and time for which it is natural to model these two characteristics of the process jointly (e.g. disease spread, species distribution). Observations of these processes are often measured as point-referenced locations in time, i.e. space-time point patterns. The difficulty in modeling these data is the complexity of the likelihood, because it needs to be integrated over the spatial and temporal domains. Krainski (2018) developed a non-separable stochastic partial differential approach (SPDE) that can be used to consider the realistic space-time evolution of the process in which the spatial and temporal autocorrelation in the latent field are linked.

The flexibility of the SPDE approach has great potential to provide novel insights through the analysis of complex spatiotemporal data. Spatiotemporal models capable of capturing local infection dynamics can generate fundamental insights into disease transmission processes and can be used to develop targeted control programmes. However, there are two main barriers preventing the wider use of this approach. First, the software is challenging to use for novices. Second, the statistical complexity of the SPDE approach is off-putting to many non-statisticians.

Using fox rabies as a case study, we illustrate how this approach can be used to capture the spatiotemporal nature of disease transmission. We model the fox rabies case data as arising from a log-Gaussian Cox point process and fit the model using R-INLA. We discuss several practical considerations for users to consider in model construction including mesh generation, model implementation, and model checking.

Univariate and Multivariate Directed Acyclic Graphical Models for Disease Mapping

Sudipto Banerjee
University of California, Los Angeles

Hierarchical models for regionally aggregated disease incidence data commonly involve region specific latent random effects that are modelled jointly as having a multivariate Gaussian distribution. The covariance or precision matrix incorporates the spatial dependence between the regions. Common choices for the precision matrix include the widely used intrinsic conditional autoregressive model, which is singular, and its nonsingular extension which lacks interpretability. We propose a new parametric model for the precision matrix based on a directed acyclic graph representation of the spatial dependence. Our model guarantees positive definiteness and, hence, in addition to being a valid prior for regional spatially correlated random effects, can also directly model the outcome from dependent data like images and networks. Theoretical and empirical results demonstrate the interpretability of parameters in our model. Our precision matrix is sparse and the model is highly scalable for large datasets. We also derive a novel order-free version which remedies the dependence of directed acyclic graphs on the ordering of the regions by averaging over all possible orderings. The resulting precision matrix is available in closed form. We demonstrate the superior performance of our models over competing models using simulation experiments and a public health application.

A longitudinal analysis of police-recorded crime and mental health in Scotland

Gergő Baranyi
University of Edinburgh

Although emerging evidence indicates a spatial patterning to psychiatric disorders, underlying mechanism leading from area deprivation to health outcomes are less understood and longitudinal studies are lacking. The aim of this study was to model long-term exposure to crime in residential areas across Scotland, as a function of deprivation, and explore its effect on mental health problems. Data were drawn from the Scottish Longitudinal Study (SLS), a 5.3% nationally representative sample of the Scottish population, including over 150,000 adults (age 16+) present at the 2001 and 2011 censuses. Information on admissions to psychiatry hospitals and prescriptions for psychotropic medications were extracted from national prescription and hospital datasets held by NHS Scotland. The 2006, 2009 and 2012 releases of Scottish Index of Multiple Deprivation provided ranks of police reported crime for each of the 6505 Scottish data zones (500-1000 residents per zone) and were linked with residential postcodes between 2004 and 2012, based on GP registrations. Group-based trajectories of changing exposure to crime were calculated with finite mixture models using censored normal distribution, while logistic regression models estimated the effects of crime trajectories on mental health. Preliminary analyses identified 5 major trajectories of crime exposure (stable low, moderate and high, increasing and decreasing). Higher levels of crime were significantly associated with prescription for psychotropic medications and admission to inpatient care after adjustment for individual covariates. Crime is an important area based determinant of mental health. Targeting high crime areas should be prioritised in order to build safer, sustainable and healthier communities.

Accounting for survey design in Bayesian disaggregation of survey-based areal estimates of proportions

`Marco H. Benedetti`
`University of Michigan`

Understanding the effects of neighbourhood factors on health requires data on characteristics of the neighbourhoods in which subjects live. However, estimates of these characteristics are often aggregated over space and time in a fashion that diminishes their utility. Take, for example, estimates derived from the American Community Survey (ACS), in which estimates for small municipal areas (e.g. census tracts) are aggregated over 5-year time periods, while 1-year estimates are only available for municipal areas with populations >65,000 (e.g. counties). Researchers may wish to use ACS estimates in studies of population health to characterize neighbourhood-level exposures. However, 5-year estimates may not properly characterize temporal changes or align temporally with other data in the study, while the coarse spatial resolution of the 1-year estimates diminishes their utility in characterizing neighbourhood exposure. To circumvent this issue, we propose a modelling framework to disaggregate estimates of proportions derived from sampling surveys, while also accounting for the survey design effect. Simulation studies and application to ACS estimates of poverty and race demonstrate our model's utility in disaggregating these estimates to a fine spatio-temporal resolution.

Spatiotemporal geostatistical preferential sampling modelling for exposure assessment in epidemiological cohort study on air pollution and maternal health

`Annibale Biggeri`
`University of Florence`

In the context of the INSIDE project - environmental epigenetics, gestational health and air pollution (PM10, PM2.5), we develop a spatio-temporal preferential sampling geostatistical model to predict exposure at each residential address of the enrolled women (see Iodice et al. 2018 Oxidative Medicine and Cellular Longevity).

199 healthy pregnant women at the “Clinica Mangiagalli”, Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico, Milan, Italy, were enrolled in the period between June 2014 and October 2015. Several fetal growth measures and blood samples were taken at the 11th week of pregnancy, as part of the prenatal screening test. A questionnaire was also obtained on relevant potential confounders. Data on PM_{10} and $PM_{2.5}$ were provided by Lombardy’s Regional Environmental Protection Agency (ARPA) Air Quality Network. In a previous paper we developed a Bayesian preferential sampling geostatistical model (Cecconi et al. 2016 Stat Methods Med Res) for PM10 exposure assessment. In this work we extend the previous model to a spatiotemporal setting and individual address location. Working on such high resolution spatial framework we modify our model in order to: 1- include information from remote sensing at fine resolution; 2- integrate misaligned information from deterministic model outputs at rougher resolution. In this example, our main interest is not anymore in showing the effect of preferential sampling on the predictions’ standard deviations but in improving accuracy of individual spatiotemporal predictions.

Spatial variations and differences in incidence rates from 1986 to 2015 in atrial fibrillation and atrial flutter in Denmark

Kristine Bihrmann[1] and Kirstine Wodschow[2]
[1]University of Southern Denmark
[2]National Institute of Public Health, Denmark

Background
Individual risk factors of atrial fibrillation and atrial flutter (AFF) have been studied, but less is known about possible neighborhood risk factors. The study objectives were to 1) examine changes in yearly AFF incidence rates (IRs) from 1985 to 2015 stratified on sex, age and socio-economic position, 2) analyse spatial variation in AFF in 2015 (latest most complete data year across all registers).

Methods
This study is based on nationwide register data of adults (>29 years) in Denmark. Yearly IRs (1987-2015) were calculated, stratified on sex, age, income and education. Spatial scan statistic in SatScan was used to detect significant clusters with high risk of AFF across age groups of 30-59, 60-69, 70-79 and >=80 years.

Results
In 2015, 20,570 incident AFF cases were observed. Incidence rate has increased from 188/100.000 person years in 1987 to 593/100.000 person years in 2015. Within the same period, the IR ratio between the lowest versus highest income quintile has increased from 1.32 to 1.54 with 95% confidence intervals [1.22,1.44] and [1.47, 1.61], respectively. One statistically significant cluster with a relative risk=1.37 (5% population, p=0.027) for age >80 years was detected.

Conclusion
The highest IRs of AFF was found with increasing age, among men and those with lowest socio-economic position. Only one significant cluster of AFF was found in 2015 in the oldest population group. In future studies of possible neighbourhood risk factors, it is recommended to adjust the spatial analysis for socio-economic position.

Mapping the burden of dengue in Indonesia to guide future Wolbachia programmes

Oliver J Brady
London School of Hygiene & Tropical Medicine, London

Background
Wolbachia infected mosquitoes are a novel vector control method that reduces dengue virus transmission. City-wide releases in Yogyakarta city, Indonesia are showing promising entomological results. Health authorities are currently considering further Wolbachia programmes and want to know in which locations such programmes would have the greatest benefit.

Methods
Here we combine multiple types of dengue transmission data and multiple mapping methods to create ensemble predictions of the spatial burden of dengue in Indonesia using generalised boosted regression models. We then combined these burden estimates with a mathematical (agent-based) model to predict the reduction in dengue burden of Wolbachia programmes in different parts of Indonesia.

Results
We estimate that 7.5 million (95% uncertainty interval [UI] 1.3-16.5 million) symptomatic dengue cases occurred in Indonesia in 2015 and were associated with 326,871 (UI 79,817 – 717,476) lost Disability Adjusted Life Years (DALYs). We map this burden at a high spatial resolution (5km x 5km), revealing high concentration in big cities with 90% of dengue cases occurring in 22.2% of land area. Implementing a Wolbachia programme targeted to these transmission focii was estimated to avert 43% (UI 23-60%) of cases over a long-term average.

Conclusions
These results suggest interventions targeted to the highest burden cities can have a disproportionate impact on dengue burden. Area-wide interventions, such as Wolbachia, that are deployed based on area covered could protect people more efficiently than individual-based interventions, such as drugs or vaccines, particularly when targeted using disease mapping approaches.

Exploring spatial patterns of greenspace use derived from GPS data from a mobile phone app

Paul Brindley
University of Sheffield

The health benefits associated with greenspace are widely recognised. There is, however, an appreciation of the need to move beyond exploring spatial patterns of greenspace provision and focus instead on the extent to which greenspaces are used. Just because greenspace exists doesn't mean people universally will use it and therefore benefit from it. Furthermore, people may not necessarily use the greenspace closest to their homes and complex relationships may exist whereby use is associated with people's activity spaces.

The development of mobile phone apps facilitate the collection of user's spatial footprints through GPS data and provides new opportunities to explore the complex patterns of people's greenspace usage. Traditional survey approaches are frequently reliant on relatively modest sample sizes and critically are costly to undertake. This work is based on GPS data collected through the mobile phone app Shmapped where data from 2,309 participants within the city of Sheffield, UK were explored.

This paper uses GIS to investigate the extent to which people are using greenspace and considers if patters vary for different user groups (for example with age or levels of deprivation). Comparisons are made against traditional survey approaches. GPS data extracted from mobile phone apps are not without their own set of limitations which are also evaluated.

New cost effective approaches of monitoring and evaluating the use of greenspace are required in order to move the greenspace agenda forward beyond simple measures of provision and towards a more nuanced analysis of people's use of greenspace.

Inequalities in mortality in Scotland: what happens when we change the level of geography used to assess area deprivation?

Denise Brown
University of Glasgow

Carstairs deprivation scores measure material disadvantage and are derived by combining four census variables. They were first created in 1981, at the postcode sector (PS) level of geography, with scores updated decennially following each census. In 2011, for the first time, Carstairs deprivation scores were calculated for Scottish PS's, data zones (DZ) and output areas (OA). Average populations in 2011 were 5 233, 759 and 114 people, respectively.

We calculate mortality rates by age, sex and area-based deprivation (at the PS, DZ and OA geographical level) for the 3-year period around the census in 2011. We also examine absolute and relative inequalities in mortality at PS, DZ and OA level, using the rate ratio of the most to least deprived areas and the slope (SII) the relative (RII) indices of inequality.

Preliminary findings suggest that inequalities in premature all-cause mortality for males and females are highest at the OA level and smallest at PS level. Generally for specific causes, the smaller the geographical level the higher the inequality. This is true, to a lesser extent, for deaths due to ischaemic heart disease, cancers and stroke and more noticeable for deaths due to drug and alcohol and chronic respiratory disease.

Latent classes for meaningful chemical mixtures analyses in epidemiology: An example using phthalate and phenol exposure biomarkers in pregnant women

Rachel Carroll
University of North Carolina at Wilmington

Latent class analysis (LCA), although minimally applied to the statistical analysis of environmental mixtures, may serve as a useful tool for identifying individuals with shared real-life profiles of chemical exposures. Knowledge of these groupings and their risk of adverse outcomes has the potential to inform targeted public health prevention strategies. This example applies LCA to identify clusters of pregnant women from a case-control study within the LIFECODES birth cohort with shared exposure patterns across a panel of urinary phthalate metabolites and parabens, and to evaluate the association between cluster membership and urinary oxidative stress biomarkers. LCA identified individuals with: “low exposure,” “low phthalates, high parabens,” “high phthalates, low parabens,” and “high exposure.” Class membership was associated with several demographic characteristics. Compared to “low exposure,” women classified as having “high exposure” have elevated urinary concentrations of the oxidative stress biomarkers 8-hydroxydeoxyguanosine (19% higher, 95% confidence interval [CI]=7%, 32%) and 8-isoprostane (31%, 95% CI=-5%, 64%). However, contrast examinations indicated that associations between oxidative stress biomarkers and “high exposure” were not statistically different from those with “high phthalates, low parabens” suggesting a minimal effect of higher paraben exposure in the presence of high phthalates. The presented example offers verification through application to an additional data set as well as a comparison to another unsupervised clustering approach, k-means clustering. LCA may be more easily implemented, more consistent, and more able to provide interpretable output. Future directions include methodological extension to temporally- and / or spatially-dependent latent classification.

Greenspace and mental wellbeing: a study of residential relocation among mid-older aged adults

Alison Carver[1], Jerome Rachele[2], Takemi Sugiyama[1] and Gavin Turrell[3]
[1]Australian Catholic University
[2]University of Melbourne
[3]RMIT University

Background
Few longitudinal studies have examined associations between exposure to greenspace and mental wellbeing as adults grow older. The potential impact on mental wellbeing of relocating to areas with more/less greenspace is unclear.
Aim
To examine temporal associations between neighbourhood greenspace and mental wellbeing among mid-older aged adults who relocated, accounting for residential self-selection.
Methods

We drew data on movers (n=1,856) from HABITAT, a multilevel longitudinal study (2007-18) conducted in Brisbane, Australia, of adults aged 40-65 years at baseline/Wave 1 (2007). Participants self-reported mental wellbeing at Waves 2-5 (2009; 2011; 2013; 2016) using the shortened Warwick-Edinburgh Mental Wellbeing Scale. Principal Components Analysis of survey-reported reasons for relocation identified four factors, proximity to: nature/greenspace; work/city; destinations; childcare/schools. At each wave greenspace/park area within 1km (Euclidean buffer) of participants' homes was objectively measured using geospatial data. To examine the association between change in exposure to greenspace and change in mental wellbeing we performed a linear fixed effects model. We adjusted for potential time-varying confounders (e.g. occupation, household income, residential self-selection), and clustering at the neighbourhood level.

Results

On average, greenspace area increased (Wave 2, 0.23 (SD 0.19) km^2; Wave 5, 0.24 (SD 0.20) km^2) and mental wellbeing increased (Wave 2, 23.3 (SD 3.5); Wave 5, 24.0 (SD 3.9)). However, there was no significant difference when comparing mental wellbeing by exposure to different amounts of greenspace following residential relocation.

Conclusion

Further longitudinal studies are required to understand whether causal relations exist between greenspace exposure and mental wellbeing among adults as they grow older.

Reconciling health facility case counts of malarial fevers with parasite prevalence survey data via Bayesian hierarchical modelling

Ewan Cameron
University of Oxford

In low resource settings with endemic malaria transmission, large-scale national parasite prevalence surveys (e.g. MIS/DHS) offer a core benchmark for mapping transmission intensity and monitoring long-term trends. However, with national surveys being logistically and financially feasible at most once per year, there is great interest in leveraging the increasingly reliable data reported by health facilities on weekly or monthly counts of clinical malaria cases (e.g. DHIS2) to allow these model-based transmission maps to reflect spatially-resolved seasonality profiles and current trends, such as might arise from the impacts of ongoing intervention programs. In this talk I will describe a suite of statistical analysis strategies developed at the Malaria Atlas Project to overcome the challenges of synthesising these two metrics, measured at different spatial scales (point data vs area-averaged data) and temporal cadences, and differing through a number of confounders including variable rates of treatment seeking and background fevers.

Detecting time-series horizontal and vertical building density at neighbourhood scales with open access remote sensing data

Tzu-Hsin Karen Chen
Aarhus University

City compaction is argued to be the most sustainable strategy for the development of cities by the European Commission's Green Paper and the UN-supported Millennium Ecosystem Assessment. However, the health effect of city compaction is dubious based on the literature because temporal profiles of urban density are usually lacking. Here, we propose an innovative remote-sensing approach to trace the two and three dimensional urban density over the past 20 years. To overcome the common dilemma in remote sensing data between temporal coverage and spatial resolution, we combine optical USGS Landsat 8 OLI and very high-resolution satellite imagery available via Google EarthTM mapping service to train machine learning models for urban density mapping. We reconstructed building density change from 1999 to 2018 using the examples of Copenhagen, Denmark and Taipei, Taiwan—large cities in Europe and Asia, to compare the performance of the developed method in different geographical settings. We define neighbourhood-level horizontal density as the ratio of the built-up area to the total area of a neighbourhood, and vertical density is a category of the mean building height (<4m, 4-10m, 10-20m, >20m) of a neighbourhood. During the session we expect to present preliminary results of our method on deriving the spatial-temporal information of building density. We expect this study will of high interest to a GEOMED audience for a glimpse at the temporal dimension of urban environments, with potential to help better understand human-environment interaction that is critical to public health solutions.

The great recession, austerity and mental health in Scotland: moderating roles of green space and social cohesion

Mark Cherrie
University of Edinburgh

The Great Recession of 2008/2009 increased unemployment and reduced living standards, which led to stress and unhealthy behaviours at a community level, due to precarious financial and social circumstances. There is little evidence on how individual circumstances (unemployment) and fiscal policy (austerity) mediated the impact of the recession on mental health, and whether the social and physical characteristics of places led to resilience. We used data on employment figures for 2004-2014 from NOMIS, estimated on the financial impact of austerity for 2010-2015 (Beatty and Fothersgill, 2015), satellite derived measures of neighbourhood greenspace for 2011 (normalised difference vegetation index - NDVI) and a neighbourhood social fragmentation index derived from the 2011 census variables (number of unmarried persons, single person households, number of privately rented households and mobility in the previous year). These were linked to over

100,000 participants of the Scottish Longitudinal Study (https://sls.lscs.ac.uk/). We developed a structural equation model to understand the impact of the recession on mental illness, derived from information on hospital admissions (2001-2011), self-reported status (2011) and antidepressant prescriptions (2009-2014). Findings indicate that only around 2% of the effect of the recession was mediated through employment status, compared with 29% through austerity measures. There was a significant interaction between employment levels and greenspace; individuals with the lowest availability of greenspace, living in areas that were most affected by the recession had the strongest positive association with mental illness. The impact of recession through austerity measures were found to be greatest in areas with the highest social fragmentation.

The effect of neighbourhood crime and deprivation on pregnancy outcomes. A quasi-experimental study using routine administrative records

Tom Clemens
University of Edinburgh

We know that health outcomes vary spatially and that characteristics of neighbourhoods, such as the level of socio-economic derivation, are strongly associated with health. Whether these effects are due to differences in terms of population composition or whether there exist independent “contextual” effects of neighbourhoods has long been debated by Geographers. The policy implications of this debate remain important; do we focus policy on the area itself or the people in the area? Part of this debate concerns the empirical challenge of isolating one explanation from the other. Much of the available empirical evidence comes from cross-sectional observational studies. These studies compare different people between different areas where it is difficult to control for confounding resulting from underlying differences between populations. Administrative health data (i.e. data collected as part of routine healthcare) provides some opportunities to tackle this problem given the full population coverage and the repeated longitudinal collection of records over time. In this presentation, I will introduce some recent work where we have used administrative records to improve our understanding of both context and composition effects in the relationship between neighbourhood characteristics and pregnancy outcomes.

Mapping environmental receptivity to dengue and chikungunya

Claudia Torres Codeço, Daniel Villela and Flavio Coelho
Fundação Oswaldo Cruz

In the vector borne disease literature, ’receptivity’ refers to environmental conditions that favor the transmission in a specific location and time. Climate conditions, particularly temperature and humidity, affect the receptivity of an area via effects on life history parameters of vectors as well as the pathogen-vector interaction. For example, large outbreaks depend on having periods of time with adequate temperatures long enough to complete generations of disease transmission. Since the generation interval of a vector borne disease shrinks when temperature increases, more generations of transmission can fit in the same season, resulting in larger epidemics. These phenomena introduce nonlinear associations between temperature and disease incidence. In this work, we calculated the temporal patterns of both dengue and chikungunya receptivity in Brazilian cities from 2010 to 2018 using climate data. A measurement of receptivity was defined as the occurrence of climate conditions sufficient for at least two generations of disease transmission. This methodology included the parameterization of probability distributions of the generation interval for both dengue and chikungunya, and the estimation of the generation interval distribution for each city as a function of temperature. The result is a dynamic map of receptivity for dengue and chikungunya transmission in Brazil. To validate this map, we assessed the temporal association between receptivity and transmission (reproductive number > 1) in 790 cities distributed along different latitudes and climate regimes. The map is integrated into Infodengue, a warning system implemented for public health response.

A spatio-temporal hierarchical Markov switching model for the early detection of influenza outbreaks

David Conesa[1], R. Amorós[2], A. López-Quílez[1] and M. A. Martínez Beneito[3]
[1]University of Valencia
[2]University of Edinburgh
[3]Public Health Research Center of Valencia

Rapidly detecting the beginning of influenza outbreaks helps health authorities to reduce their impact. Accounting for the spatial distribution of the data can greatly improve the performance of an outbreak detection method by promptly detecting the first foci of infection. The use of Hidden Markov chains in temporal models have shown to be great tools for classifying the epidemic or endemic state of influenza data, as a result their use in spatio-temporal models for outbreak detection could be so revealing. In this work, we present a spatio-temporal Bayesian Markov switching model over the differentiated incidence rates for the rapid detection of influenza outbreaks. This model focuses its attention in the incidence variations to better detect the higher increases of early epidemic rates even when the rates themselves are relatively low. The differentiated rates are modelled by a Gaussian distribution with different mean and variance according to the epidemic or endemic state. A temporal autoregressive term and a spatial conditional autoregressive model are added to capture the spatio-temporal structure of the epidemic mean. The proposed model has been tested over the USA Google Flu Trends database to assess the relevance of the whole structure.

An epidemiological risk assessment for arsenic in private water supplies based on geology

Helen Crabbe[1], Tony Fletcher[1,2], Michael Watts[3], Rebecca Close[1], and Giovanni Leonardi[1,2]
[1] Public Health England

[2]London School of Hygiene and Tropical Medicine
[3]British Geological Survey

Introduction
About one million people in the UK are served by private water supplies (PWS) where mains connection is not possible. Chronic exposure to chemicals in PWS may have an adverse effect on health. South West England is an area with elevated arsenic concentrations in groundwater supplying over 9000 domestic property PWS. There remains uncertainty as to the extent of the population exposed to arsenic, and factors predicting such exposure. We developed a risk assessment model based on simplified geology to predict exposure to arsenic in PWS.
Methods
We invited known PWS users on a register in Cornwall to take part in a water sampling programme from 2011-13. PWS were sampled by random selection within strata defined by a simplified geological classification. We developed a model using the UK Radon programme methodology to predict the distribution of arsenic concentrations for PWS on similar geologies where no measurements were available. Using the distribution per geology, we predict the proportion of properties that would be affected by high concentrations.
Results
508 domestic dwellings had their PWS sampled. Within each simplified geological area, concentrations were found to have log-normal distributions. Across the same areas, the proportion of dwellings predicted to have drinking water over the prescribed concentration value (PCV) for arsenic ranged from 1-20%. From these results, a pilot predictive model was tested calculating the number of PWS above the PCV for arsenic.
Conclusions
A tool is under development to help local authorities predict the number of dwellings that might fail the PCV for arsenic, based on bedrock geology. This could help prioritize areas for testing and estimate size of resources needed for regulation purposes. However, further work is essential to validate the hypothesis and improve the predictive capability through groundtruthing (sampling versus model results) across a range of geological classifications.

Measuring the Access to Healthy Assets and Hazards for Great Britain

Konstantinos Daras
University of Liverpool

Our study details the creation of the second version of the Access to Healthy Assets and Hazards index for Great Britain. This composite index derived from a series of national open source low-level geographical measures of accessibility to health-related features of neighbourhoods in Great Britain. We create 16 measures across four domains: retail environment (fast food outlets, gambling outlets, pubs/bars/nightclubs, off-licences, tobacconists), health services (General Practitioners, pharmacies, dentists, hospitals, leisure centres), physical environment (blue & green spaces) and the air quality (SO2,NO2, PM10). We find the most accessible healthy areas are concentrated in the periphery of the urban cores, whilst the least accessible healthy areas are located in the urban cores and the rural areas. All data including the index, domain scores and input values are available to downloaded freely via the CDRC interactive web mapping tool.

Identifying spatial dynamics of infectious disease spread via machine learning classifiers

Rob Deardon
University of Calgary

Information obtained from statistical infectious disease transmission models can be used to inform the development of containment strategies. Inference procedures such as Bayesian Markov chain Monte Carlo are typically used to estimate parameters of such models, but are often computationally expensive. However, in an emerging epidemic, stakeholders must move quickly to contain spread. Here, we explore machine learning methods for carrying out fast inference via supervised classification, with a focus on deep learning. We apply our methods to simulated epidemics through two populations of swine farms in Iowa, and find that the random forest performs well on the denser population, but is outperformed by a recurrent neural net model on the sparser population.

Spatio-temporal modeling of neighborhood level risks for dengue, chikungunya, and Zika in Cali, Colombia

Michael R. Desjardins
University of North Carolina at Charlotte

Vector-borne diseases (VBDs) affect more than 1 billion people a year worldwide, cause over 1 million deaths, and cost hundreds of billions of dollars in societal costs. Mosquitoes are the most common vectors, responsible for transmitting a variety of arboviruses. Recently, several particular VBDs, most notably dengue fever, chikungunya, and Zika have been responsible for the majority of the burden caused by mosquito-borne diseases. These three diseases are primarily transmitted by the Aedes Aegypti and Aedes Albopictus. Since both Aedes species are peridomestic, container-breeding mosquitoes, vector surveillance should begin at the neighborhood level - where a variety of local factors may increase the risk of transmission. Dengue has been endemic in Colombia for decades, while chikungunya, and Zika first appeared in 2013. For this study, we examine weekly cases of three VBDs in Cali, Colombia from 2015-2016. Space-time conditional autoregressive models were developed to quantify how disease risk is influenced by socioeconomic, environmental, and accessibility risk factors, and monitoring progression of disease risks over time. Our model is also capable of identifying regions with high risk clusters at neighborhood-level. Statistical inference is drawn under Bayesian paradigm using Markov Chain Monte Carlo techniques. The results can provide detailed insight about the spatial heterogeneity of disease risk and the associated risk factors at a fine-level, while informing public health officials and community leaders to motivate at-risk neighborhoods to take an active role in vector surveillance and control, while improving educational and surveillance resources throughout the city of Cali.

Exposure to air pollution in early life and cognitive ability: a prospective 1936 birth cohort

Chris Dibben
University of Edinburgh

There is increasing evidence that air pollution is linked to the development of cognitive ability in early life. In the UK levels of pollution are relatively low compared to other parts of the world. However exposure in the 1930's was considerably higher. Making use of the unique measurement of cognitive ability for a majority of 11 year old in Scotland and its link to the Scottish Longitudinal Study (SLS) we explore the effect of exposure of high levels of air pollution in childhood on cognitive ability. We model air pollution using an EMEP4UK chemical transport model and residential density, estimated using an image analysis algorithm run over a digitised historical map series, to produce detailed spatial estimates of particulate matter (PM) for the year 1935. Controlling for other predictors of cognitive ability we then estimate the additional effect of residential air pollution on cognitive ability and then the impact of that difference on highest educational achievement and adult social class position. Understanding the lifetime effect of the levels of air pollution being experienced by child in many cities around the world is of clear importance.

Urbanization and health in Latin America: the SALURBAL Study

Ana V. Diez Roux
Drexel University

Urbanization rates are increasing all over the world. Understanding the drivers of health in cities and the most effective policies to promote urban health and health equity is therefore an important need globally. This presentation will present the rationale for, the aims, the approach and preliminary results from an innovative partnership to study urban health in Latin America: the Salud Urbana en America Latina (SALURBAL) project. This multicountry partnership brings together observational studies, policy evaluation, systems modelling and policy outreach to generate and disseminate policy relevant evidence on urban health applicable worldwide.

Developing a locally adaptive spatial multi-level logistic model to analyse ecological effects on health

Guanpeng (Gavin) Dong
University of Liverpool

Geographical variable distributions often exhibit both macro-scale geographic smoothness—implying positive spatial dependency—and micro-scale discontinuities — suggesting implicit local step changes. Nonetheless, accounting for both effects in a unified statistical model poses great challenges, especially when data under study is associated with a multi-scale structure and the response variable is non-Gaussian.

We develops a locally adaptive spatial multi-level logistic model to examine binomial response variables, which integrates an innovative locally adaptive spatial econometric model with a multi-level model. It takes into account global spatial auto-correlation, local step changes, and vertical dependence effects arising from the multi-scale data structure. Another appealing feature is that the spatial correlation structure, implied by a spatial weights matrix, are learned along with other model parameters via an iterative estimation algorithm, rather than being presumed to be invariant. Bayesian Markov chain Monte Carlo (MCMC) samplers are derived to implement this new spatial multi-level logistic model. A data augmentation approach, drawing upon recently devised Pólya-Gamma distributions, is adopted to reduce computational burdens of calculating binomial likelihoods with a logit link function. The efficacy of the developed model is evaluated by a set of simulation experiments, before being applied to analyse self-rated health for the elderly in Shijiazhuang, the capital city of Hebei province, China. We utilise a unique large-scale individual census record data, augmented with the finest-resolution census geography publicly available in China. Model estimation results highlight a nuanced geography of self-rated health, and identify a range of individual- and area-level correlates of health for the elderly.

The Australian Cancer Atlas: mapping reliable small-area estimates of cancer incidence and survival

Earl Duncan and Erin Peterson
Queensland University of Technology

The Australian Cancer Atlas is a world-first innovation that combines cutting edge digital technology and sophisticated statistical models to map patterns in cancer burden at the small-area level. The atlas includes estimates of cancer incidence and survival for 20 of the most common cancer types, for males, females, and persons combined, for 2148 areas. It is freely accessible online, visually appealing, interactive, and informative – a useful tool for cancer researchers, policy makers, the media, and the general public. In this talk, I give a brief history of the development of the atlas, discuss the technical details of the spatial models, the importance of spatial smoothing, several novel approaches to visualising uncertainty on the atlas, and other specifically-designed features of the atlas. I also discuss some of the major hurdles our teams faced, in terms of both modelling and visualisation, the impact of the atlas in Australia and internationally, and directions for the future.

Association between low-level arsenic concentration in drinking water and incidence rate of stroke

Annette Ersbøll
University of Southern Denmark

Humans are exposed to arsenic through ingestion of water and food. Arsenic in drinking water is present in the inorganic form, which is associated with several chronic health consequences, thus representing a threat to human health. The guideline value for arsenic in drinking water in Denmark is 5 µg/L at the consumers tap. The aim of the study was to investigate if arsenic in drinking water, even at low concentration, was associated with higher incidence rate of stroke in Denmark.

In the Danish Diet, Cancer, and Health cohort of 57,053 people aged 50-64 years at enrolment in 1993-1997, we identified 2,195 cases of incident stroke in the Danish National Patient Register between enrolment and November 2009. Associations between arsenic concentration and incidence rate of stroke were analyzed using a generalized linear model with a Poisson distribution. Incidence rate ratios (IRR) were adjusted for differences in age, sex, calendar-year, lifestyle factors, and educational level.

At baseline addresses, the arsenic concentration in drinking water ranged from 0.03 to 25 µg/L. The adjusted IRRs were 1.17 (95% CI: 1.04-1.32) for the highest arsenic quartile (1.93-25.3 µg/L) when compared with the lowest quartile (0.049-0.57 µg/L). Having ever (versus never) lived at an address with 10 µg/L or more arsenic in drinking water resulted in an IRR at 1.44 (95% CI: 1.00-2.08) for all strokes and 1.63 (95% CI: 1.11-2.39) for ischemic strokes. The results indicate that even low-level arsenic concentration in drinking water is associated with higher incidence rate of stroke.

Does survival after acute myocardial infarction depend on where you live?

Annette Ersbøll[1] and Thora Kjærulff[2]
[1]University of Southern Denmark
[2]National Institute of Public Health, Denmark

Objective
The aim of this study was to describe the interplay between geographical and social inequalities in survival after incident acute myocardial infarction (AMI) and examine whether geographical variation in survival exists when accounting for individual-level and neighbourhood-level sociodemographic characteristics.

Methods
Ringmap visualization and generalized linear models were performed to study post-AMI mortality. Three individual-level analyses were conducted: immediate case fatality, survival among patients alive one day after admission and survival among patients alive 28 days after admission. Generalized linear models with a geographically structured component as a random effect was performed in R by use of the Bayesian Integrated Nested Laplace Approximations (INLA) method.

Results
In total, 99,013 incident AMI cases were registered between 2005 and 2014 in Denmark. Survival after AMI tended to correlate with sociodemographic indicators at the municipality level. In individual-level models, marked geographical variation in immediate case fatality was observed with residual odds ratios of AMI death ranging from 0.70 to 2.96 across the municipalities after accounting for sociodemographic characteristics. High mortality was observed mainly in northern parts of Jutland. Less geographical variation was observed in survival among patients alive one day (residual mortality rate ratios (MRRs) ranging from 0.88 to 1.24) and no clinically relevant variation in MRs was observed among patients alive 28 days after AMI (residual MRRs ranging from 0.92 to 1.02).

Conclusion
Marked geographical inequality was observed in immediate case fatality. In contrast, no geographically unequal distribution of survival was found among patients who survived 28 days after AMI.

Using the NZ Index of Multiple Deprivation (IMD) to inform research, policy and populations

Daniel Exeter
University of Auckland

For the past 20 years, the New Zealand Index of Deprivation (NZDep) has been widely used to measure area-level deprivation in NZ. We recently developed the New Zealand Index of Multiple Deprivation (IMD), a set of tools for identifying concentrations of deprivation in New Zealand. The IMD measures deprivation at the neighbourhood-level in custom designed data zones that have an average population of 712. Data zones are aggregations of census meshblocks (approximately 8 meshblocks per data zone). We identified 28 indicators of deprivation, categorised into 7 Domains of deprivation (Employment, Income, Crime, Housing, Health, Education, Access) which can be used separately or combined as the IMD. Since its introduction, the IMD has been used throughout the country by researchers, non-government organisations, as well as local and central-government departments. Following a brief introduction to the landscape of deprivation according to the IMD, we compare and contrast the IMD and NZDep. Next, we showcase some examples of how the IMD and data zones have been used in different settings to provide a more nuanced understanding of area deprivation in New Zealand.

Mapping 'Rattiness' to understand leptospirosis transmission in Salvador, Brazil: A geostatistical modelling approach to combining three rat metrics

Max Eyre
Lancaster University

Leptospirosis is a zoonotic disease which causes over a million human cases annually, the majority of which occur in vulnerable urban communities in tropical and subtropical regions. Rats are the main reservoir for infection and contaminate the local environment with their urine. Despite this, little is known about the relative contribution of rat population dynamics in transmission. Assessing rat population dynamics is complicated by a lack of a gold-standard measurement instrument and an inability to differentiate between rat presence, abundance and activity in some instruments. In this study three measures of rat population dynamics were used: rat trapping, track plates and identification of rat signs. It was assumed that these three outcomes are proxies for a single latent process of 'Rattiness'.

Data was collected in four communities in Salvador, Brazil in 2018. Covariate data was collected from household survey questions (n=951), ecological surveys, mapping of sewers and refuse in the field and remote-sensing derived data. The association between environmental and household-level risk factors and Rattiness was evaluated using a multivariate generalized linear geostatistical model. Relative Rattiness was then predicted across the four study areas and compared to leptospiral infection risk maps.

This study demonstrates how multiple imperfect metrics can be combined within a geostatistical framework in an eco-epidemiological context. Evaluating the relative contributions of the three rat metrics to Rattiness, and their associations with environmental and household-level risk factors is an important first step in understanding the role of rodent populations in leptospirosis transmission.

Bayesian spatial smoothing methods for analysis of spatial survey data

Christel Faes
Hasselt University

Bayesian spatial smoothing methods play an important role in the field of spatial epidemiology and small area estimation. Health surveys are often characterised by a complex design, with stratification and unequal sampling weights. As a result, the use of design weights is indispensable in the estimation process. Recently, several spatial smoothing models have been investigated and compared to obtain reliable estimates of the spatial trend. Traditional design-based estimates (such as the Horvitz-Thompson estimator) are compared with several model-based approaches and a spatial predictive approach. The model-based approaches, extensions of the Fay-Herriot model, are based on area-level data. The predictive approach uses unit-level data.

Another feature of health surveys is the presence of missing information. The impact of individual nonresponse, which is the result of a selected individual not responding to a question, on the models is investigated, and extensions of the models are proposed. Further, as in many surveys only a subset of areas is included in the survey, also the number of selected areas is investigated. Here, we can make a distinction between in-sample and off-sample areas, in which the estimation of the off-sample areas is based on the borrowing of information from the available in-sample areas. It is investigated under which circumstances the underlying spatial trend can be resolved. Methods are illustrated using different surveys: the Belgian health interview survey, the Florida BRFSS survey and the poverty and social impact analysis survey in Mozambique.

Ecological inference for infectious disease data, with application to vaccination strategies

Leigh Fisher
Fred Hutchinson Cancer Research Center

Disease surveillance systems provide a rich source of data regarding infectious diseases, aggregated across geographical regions. The analysis of such ecological data is fraught with difficulties, and unless care and suitable data summaries are available, will lead to biased estimates of individual-level parameters. We consider using surveillance data to study the impacts of vaccination. To catalogue the problems of ecological inference we start with an individual-level model, that contains familiar parameters, and derive an ecologically consistent model for infectious diseases in partially vaccinated populations. We compare with other popular model classes and highlight deficiencies. We explore the properties of the new model through simulation, and demonstrate that under standard assumptions, the ecological vaccine model provides less biased estimates. We then apply this approach to data collected on measles outbreaks in Germany from 2005-2007.

Environmental change and socio-political drivers of malaria risk in El Oro province, Ecuador

Isabel Fletcher
London School of Hygiene & Tropical Medicine

Vector-borne disease risk is determined by a multitude of factors, including climate and socioeconomic conditions. Understanding the drivers of spatiotemporal variation in disease risk can be used to inform future elimination efforts. Environmental change, including climate and land-use, has a well-established relationship with vector-borne diseases and can be used to help develop predictive models about how future risk may change. Malaria is resurging in areas of the world that had previously suppressed transmission, which is threatening global elimination efforts.

El Oro province, Ecuador eliminated malaria in 2011 but recently, along with other countries in the region reported an increase in malaria, which is likely due to imported cases from Venezuela. A Bayesian hierarchical mixed model was formulated to investigate the drivers of malaria incidence in El Oro between 1990-2010 to understand the factors that contributed to disease risk and may increase the risk of resurgence. The model included climate covariates and land use change. Spatial and temporal random effects were included in the model to account for additional unknown and unmeasured factors, such as differences in intervention practices and political instability, along with spatial dependency structures.

Maximum temperature was a significant predictor of malaria incidence, with higher temperatures associated with an increase in malaria risk. The proportion of urbanised areas also have a positive and statistically significant association with malaria risk in El Oro. These results can be built upon to inform future models of disease risk to anticipate how risk may change in El Oro given regional resurgence threats.

Disease mapping using complex survey data

Geir-Arne Fuglstad
Norwegian University of Science and Technology

Many people living in low- and middle-income countries are not covered by vital registration systems. Consequently, household sample surveys with complex designs are often used to estimate health and population indicators. The traditional approach for handling complex survey data is design-based, but breaks down at fine spatial and temporal scales. Developing better model-based methods for taking full advantage of these types of data at fine spatio-temporal scales and understanding the accuracy of such methods are essential for providing the officials in each country and the United Nations with the tools necessary for reaching the Sustainable Development Goals (SDGs) of the United Nations. In this talk, I will discuss a new fine-scale spatio-temporal model for under-five mortality in Kenya and outline how to construct a framework for assessing the predictive performance of a new model and traditional approaches through proper scoring rules.

Spatio-temporal two-stage models for environmental research

Antonio Gasparrini
London School of Hygiene and Tropical Medicine

Novel big data resources offer exceptional opportunities for environmental research, allowing linkage of health data with high-resolution exposure measurements in large populations and study areas. However, this new setting presents important analytical and computational issues, including the problem of modelling potentially complex associations varying over spatial and temporal units, the consideration of confounders and effect modifiers measured at different geographical levels, and the exceptional computational burden of performing analyses spanning entire countries and several decades.

In this contribution, we present a novel spatio-temporal two-stage design to perform small-areas analyses in environment-health epidemiological investigations. The two stages involve, respectively, the application of case time series Poisson models to estimate localized period-specific associations, and an extended meta-analytical framework to pool information across first-stage units. The first-stage regression can define complex exposure-lag-response associations and control for temporal and spatial confounders, while the second-stage meta-analysis can account for complex correlation structures of repeated estimates collected over time and space. The design can be flexibly tailored to different settings and modelling problems. This framework will be illustrated in a small-area analysis of temperature-mortality associations using data collected in 34,753 Lower Layer Super Output Areas (LSOAs) in England and Wales in the period 1981-2016. Different designs will be defined and applied to investigate geographical differences in the increased risks associated to heat and cold, to explore potential temporal variations, and to assess spatially and time-varying characteristics that can potentially modify the relationships.

Household survey on influenza-like illness and health related behaviour in urban quarters of Basel, Switzerland

Nina Goldman
University of Basel

Seasonal influenza is a preventable, often underestimated infectious disease. In Switzerland, it causes between 1,000 and 5,000 hospitalizations and about 400-1,500 deaths annually. Vaccination is the most effective strategy to reduce virus transmission and decreases the severity of influenza, number of complications and deaths. Yet, public confidence of influenza vaccines is decreasing which leads to reductions in vaccination rates thereby limiting its effectiveness. Community variation in attitudes, beliefs and behaviours about influenza vaccination have been identified as an important issue, as well as the lack of comprehensive assessment of community opinions and behaviours related to vaccination. For this study, which is part of a large interdisciplinary project funded by the Swiss National Science Foundation, we distributed 30,000 questionnaires (return rate 27.2%) in ten urban quarters of Basel, making this the largest survey on influenza undertaken so far in Switzerland. Respondents could voluntary self-disclose their location on a map, which enabled us to do a spatial analysis of the survey data. The survey covered topics on incidences and severity of illness, health related behaviours, opportunities for exposure in close contact environments, aspects of the city environment, procurement of health information and personal data. Selected results are presented to gain insights concerning (1) the geographic spread of the variables by exploring socio-spatial factors, (2) the determinants of health related behaviour and influenza-like illnesses (ILI) and (3) analyse a specific health related behaviour, namely vaccination, on an urban quarter level.

Environmental Health: A LiDAR Assessment of Urban Neighbourhoods

Sue C. Grady
Michigan State University

Lead is a cumulative human toxicant that affects in particular the neurological and vascular systems. Children are vulnerable to the uptake of lead if they live in post-industrial cities with older housing stock (lead-based paint and pipes) and/or proximity to industrial lead emissions (lead-airborne and deposition). This study will conduct an exploratory environmental lead assessment using LiDAR (light detection and ranging) optical remote sensing techniques in Detroit, Michigan and industrial-lead dispersion modeled concentrations (atmospheric and deposition) and childhood blood lead levels as validation. Airborne LiDAR collected by the State of Michigan in 2015 for the study area has multi-spectral bands at 12 inch resolution. 3-D point clouds are interpolated as raster grids for digital surface and intensity modeling by which to delineate built and natural environments and extract urban features that may comprise lead for future remediation.

Is the association between the density of fast food outlets and obesity confounded by socioeconomic context? A longitudinal analysis of children in the Millennium Cohort Study

Mark Green
University of Liverpool

Regulating the location of fast food outlets (e.g. around schools, in areas with high densities) is being used by local governments to tackle childhood obesity. While this work is supported by a large academic literature, systematic reviews demonstrate mixed associations for whether there is a relationship between the density of fast food outlets and obesity. This may be because studies fail to adequately adjust for socioeconomic context, utilising simple measures or do not correctly specific it as a confounder on this relationship (given that it is associated to both density of fast food outlets and obesity independently). Improving the quality of evidence will help inform whether regulatory interventions will be successful. This study uses data from the Millennium Cohort Study, a representative longitudinal survey following the lives of a group of children born in Great Britain in 2000 (n 10,000). Data from waves 5 (2012) and 6 (2016) are used to measure the multiple dimensions of child and parental socioeconomic context (e.g. material disadvantage, social support) and how they are associated to the consumption of fast food and anthropometric measures (e.g. body mass index, body fat). Location of shops selling fast food is derived from the Food Standards Agency for both time points and linked to participant's residential address. A latent variable modelling framework to measure socioeconomic context is used within a Structural Equation Model to test specific pathways. Models are run separately for each wave, and then cross-lagged models are used to extend the model longitudinally.

Do spatiotemporal social fragmentation and deprivation affect suicide risk? A Dutch register-based case-control study

Paulien Hagedoorn
Utrecht University

Being the fifth leading cause of death among middle-aged adults worldwide, suicide remains a key public health issue. There is increasing interest in the influence of the social living environment on suicide risk, in addition to individual characteristics. However, findings so far are mixed. A reason may be the diverse spatiotemporal ways in which the context is defined. Bespoke neighbourhoods around residential locations have not been applied in suicide research, even though these might reflect local exposure more accurately than administrative areas. This study looks into the associations between suicide mortality and social and material deprivation using address-based individualized neighbourhoods at different buffer sizes. We also assess whether this association is moderated by length of residence and individual characteristics. Using a nested case-control design, all suicides among adults from 2007 until 2016 are selected from longitudinal Dutch register data and matched to 10 random controls based on age, sex and calendar time. Deprivation

and social fragmentation over time is represented by annual indices calculated for 300, 600 and 1000 meter circular buffers around each individual's residential address using GIS. For each case and corresponding controls the individual and neighbourhood characteristics at time of suicide are selected. Preliminary results from fully adjusted conditional logistic regressions show that male suicide risk decreases with increasing social fragmentation and deprivation, while no association was found for women. Results are similar across buffer size and length of residence. These findings suggest that living in an affluent and cohesive neighbourhood does not necessarily protect against suicide.

Projection-based Methods for Fitting Spatial Generalized Linear Mixed Models

Murali Haran
Penn State University

Spatial generalized linear mixed models (SGLMMs) are popular for non-Gaussian spatial data. Inference for SGLMMs can be computationally challenging. I will describe a random projections-based approach for efficiently fitting SGLMMs. I will explain how the approach may be useful in both Bayesian and maximum likelihood settings, using Markov chain Monte Carlo (MCMC) and MCMC Expectation-Maximization algorithms respectively. This talk is based on joint work with Yawen Guan and Jaewoo Park.

Daily mortality and air quality: using multivariate time series with seasonally-varying covariances

Guowen Huang
University of Toronto

We studied the association of daily mortality with short-term variations in the ambient concentrations of PM2.5, NO2 and O3 in Vancouver and Toronto. Firstly, a multivariate time series model within Bayesian framework was proposed for exposure assessment, which was a mixture of Gamma and Half-Cauchy model, with the latter being used to capture the heavy tail of data distribution. Seasonally varying covariance among pollutants was allowed, and the pollution model was implemented in stan with parallel subthreads to speed computation. Then a case-crossover design and conditional logistic regression disease model was used to relate exposure to mortality data during 1981-2012.

Estimating the changing nature of Scotland's health inequalities using a multivariate spatio-temporal model

Eilidh Jack, Duncan Lee and Nema Dean
University of Glasgow

Health inequalities are the unfair and avoidable differences in people's health across the population between different social groups. These inequalities have a huge impact on people's lives, particularly those who live in the poorer end of the socioeconomic spectrum as they result in prolonged ill health and shorter lives. Most studies estimate health inequalities for a single disease, but this will give an incomplete picture of overall inequality. Here we propose a novel multivariate spatiotemporal model to quantify health inequalities in Scotland across 3 major diseases, which will enable us to better understand how they have changed over time. We are also interested in estimating health inequalities between Scotland's 14 regional health boards, who are responsible for the protection and improvement of their population's health. The methodology is applied to hospital admissions data for cerebrovascular disease, coronary heart disease and respiratory disease from 2003 to 2012 across Scotland.

GEM: A domain-specific modelling language for rapid development of inference and simulation for epidemic models

Chris Jewell
Lancaster University

Responding to epidemics is a grave challenge of the 21st century as new diseases emerge alongside climate, environment, and social change. Recent innovations in statistical approaches to epidemic models have provided cutting-edge decision support solutions for outbreaks. However, these demand advanced mathematics, statistics, computing, and epidemiology from the researcher, putting state-of-the-art analysis out of immediate reach of many in epidemics researchers. The GEM project aims to build a domain-specific modelling language (DSML) for rapid development and communication of epidemic models. The design specification includes an intuitive, expressive, and flexible language in which to describe disease transmisison models, cutting edge Bayesian parameter inference and simulation algorithms, and a computation layer which automatically makes use of available hardware (CPU, GPU, cluster, etc) for maximum scalability and performance. This talk will describe the approach to decomposing a general class of state transition models suitable for efficient computation, leading to the possibility of automatic code generation through the DSML.

Creating a small area deprivation index for Brazil to assess inequalities in mortality

Elzo Júnior[1], Mirjam Allik[1], Dandara Ramos[1], Maria Yury Ichihara[1], Alastair Leyland[2] and Ruth Dundas[2]
[1]Center for Data Integration and Health Knowledge
[2]University of Glasgow

Introduction
Area-based deprivation indices have been used in the UK for over 30 years, and in many other countries, such as New Zealand, Canada, France and Germany. These deprivation indices are important for understanding inequalities in general and health inequalities.

Aim
To develop an area-based deprivation index and describe inequalities in mortality for the population of Brazil.

Methods
IBGE Demographic Census Data were used to develop a deprivation index based on Census tracts, (average 400 households). Candidate indicator variables: education, literacy, income, access to clean water, sanitation infrastructure and housing type, were assessed for quality, completeness and utility; combined to create the deprivation measure. Mortality data from 2009-2013, obtained from Brazil Ministry of Health, were geocoded, assigning each death to a census tract. Age standardised mortality rates were calculated overall and by region for each decile of deprivation.

Results
Final indicator variables selected for the index are literacy, income and housing. The mean of the proportion of people across Census tracts classified as deprived within each domain are: income=16.6%, literacy=11.6%, housing=19.8%. For income and literacy, the distribution across census tracts was exponential, but this was not true for housing. There were differences in the distribution of the indicator variables across the geographic regions of Brazil, and they also varied by urban/rural indicators. Variability in geocoding deaths by region resulted in non-linear inequalities in mortality using the small-area deprivation index.

Discussion
Further work is needed to refine the inequalities in mortality analyses to account for missing geocoded deaths.

From spatial data to health action: The March of Dimes Geo-Equity App

Michael R Kramer
Emory University

Eliminating persistent racial, economic, and geographic disparities in preterm birth and infant mortality in the U.S. is a top national public health priority, and at the core of the mission for non-governmental advocacy and research foundations such as the March of Dimes. March of Dimes (MoD) were early adopters in web-based delivery of state-level perinatal data (e.g. Peristats website) but researchers, health planners, and advocates alike require more robust small-area perinatal data, and tools for understanding how social and policy context give rise to population health. To meet this need, we have a) produced county-level, race- and year-specific estimates of very, moderately, and late preterm birth at the U.S. county scale using CAR Bayesian disease mapping tools; b) assembled evidence-based social, environmental, and policy measures at the county and state scale; and c) developed an interactive Shiny web-app that permits investigation of univariate and bivariate statistical relationships (e.g. among perinatal measures or between environmental and perinatal variables) alongside visually-linked maps with service overlays to facilitate local planners' understanding of their state and regional perinatal burdens and social and health resources. In this paper we will describe the estimation, tool functionality, and present results from focus groups with MoD staff representing different use cases including national strategy planning around policy initiatives, and local advocacy and communication around sub-state hotspots and medical care deserts. Delivering robust spatial data to end-users requires attention to typical use-specific needs, interface quality, and ability to interactively understand relationships between local resources and local health burden.

Maternal and Infant Health in Central-Michigan: Environmental Investigation of Industrial Emitted Neurotoxins on Adverse Birth Outcomes

Amanda Kreuze
Michigan State University

Preliminary analyses demonstrate that the Tri-Cities region of Central Michigan, defined as eight counties within the East Central Prosperity Region, has elevated birth defects compared to all other counties in the state, excluding the Detroit Metropolitan Area. The Tri-Cities region of Central Michigan has substantial industry with over 1,660 active facilities. Among these, 28 facilities are classified as major polluting facilities by the U.S. Environmental Protection Agency. Together, these facilities emit an array of carcinogenic and neurotoxic chemicals into the air, waterways, and on land. This health geographic research will focus on neurotoxins because of their potential reproductive and teratogenic impacts on pregnancy and fetal health. Geospatial technologies will be used to conduct a neurotoxic chemical assessment, to estimate industrial emitted airborne and depositional concentrations within the Tri-Cities region of Central Michigan. These estimates will be used to define maternal exposure at place of residence, during critical windows of pregnancy health. Multilevel models will be estimated to assess the effects of maternal neurotoxin exposure on adverse birth outcomes, including birth defects, controlling for potential maternal and infant confounding variables. The findings from this research will inform environmental regulatory and public health policies and health care practice in Michigan, for translation into other industrial cities in the United States and other countries experiencing rapid industrial-related economic growth.

Spatial tracking of the current Ebola outbreak in Congo

Ashok Krishnamurthy
Mount Royal University

We present a spatial Susceptible-Vaccinated-Exposed-Infectious-Recovered-Dead (S-V-E-I-R-D) compartmental model of epidemiology to capture the transmission dynamics and the spatial spread of the ongoing Ebola outbreak in the eastern region of Kivu in Congo. For the current outbreak in Congo we use registered data (province-wide weekly counts of total Ebola cases and confirmed dead) up to June 12, 2019 from the World Health Organization (WHO) situation reports.

“Data Assimilation” is a general class of techniques for tracking a state vector in time, using Bayesian updates applied to a dynamic model. Our results for the 2013-16 West African Ebola outbreak suggest that forecasting incidence using data assimilation can be produced in the domain of quantitative tracking of an epidemic across space and time (Krishnamurthy and Cobb, GEOMED, 2015). We observed that the prediction improves as data is assimilated over time. The data assimilation layer receives sparse and error-prone epidemiological data from the field and uses this data to perform corrections to the current state vector of the epidemic. In other words, it enhances the operation of the spatial SVEIRD model by periodically executing a Bayesian correction to the state vector(s), in a way that is, at least arguably, robust and statistically optimal.

The projected number of newly infected and death cases up to August 31, 2019 are estimated and presented. We provide a discussion and interpretation of our results. The data assimilation method presented herein can be applied to a large class of compartmental or even agent-based models.

Computational approaches to Bayesian Spatio-Temporal Surveillance of small area health data

Andrew Lawson
Medical University of South Carolina

In this presentation, I will review some approaches to ST surveillance when MC sampling is used for posterior characterization. With infectious disease as the main focus, I will consider the use of a variety of metrics, which are either residual – based, or based on posterior functionals, such as SCPO, SKL, and surveillance residuals. We propose using new combinations of these metrics which include partial prediction using two stage models. We also explore the use of directional resultants in attempting to predict the spatial pathways for future infection in an infectious disease context. If time permits, I will also contrast the use of sequential MC as compared to conventional McMC when metrics are used.

Predicting tsetse abundance: remote sensing and the impact of insecticide treated cattle in Tanzania

Rachel Lea
Liverpool School of Tropical Medicine

Tsetse flies transmit trypanosomes which cause African trypanosomiasis in humans and livestock. In East Africa, wilderness areas supporting wild hosts and tsetse are ‘hotspots’ of trypanosomiasis, posing a risk to people and livestock in surrounding areas. Protecting people and livestock at the interface of wilderness and farming areas is difficult because treatment of wild hosts with drugs is impossible and controlling tsetse across wilderness areas is prohibitively expensive. A more cost-effective approach might be to focus interventions on interface areas where risk of trypanosomiasis is high.

We used remotely sensed data for Normalised Difference Vegetation Index, land surface moisture and temperature with a digital elevation model, in a geostatistical analysis of tsetse abundance. We fitted the model to data from inside the Serengeti National Park, to produce a 0.5 km resolution predictive map. The model was then tested using abundance data from farming areas at the interface of the Serengeti and Saadani National Parks.

The model correctly predicted observed declines in tsetse abundance at the interface, due in part, to anthropogenic changes in vegetation. In some farming areas, tsetse were predicted to be abundant but entomological surveys found none. Questionnaires of livestock keepers and quantification of insecticides on cattle in these areas suggested that widespread treatment of cattle with pyrethroids is responsible for declining tsetse populations in these areas.

Remotely-sensed data can be used to produce fine-scale predictive maps of trypanosomiasis risk, and interventions by livestock keepers to control tsetse in interface areas can contain the spread of trypanosomiasis.

Estimating health inequalities from multiple sets of spatially misaligned disease counts of differing severities

Duncan Lee, Craig Anderson, Nema Dean, Eilidh Jack and Kamol Sanittham
University of Glasgow

Spatio-temporal patterns in population-level disease risk are often estimated from data relating to a set of irregularly shaped areal units, such as electoral wards or census tracts. The resulting inference is conditional on the choice of units, and will change if the study region is partitioned into a different set of areal units. This is known as the modifiable areal unit problem (MAUP), which is a criticism of this type of spatial modelling. Another major problem is how to obtain comparable inference if the set of areal units change in the data set being modelled, such as the boundaries being redrawn between two different years of data, or two data sets with different spatial supports. This talk proposes a statistical framework for solving these problems, by using the areal unit data to obtain inference on the spatio-temporal pattern in disease risk on a regular grid. The pixels in this regular grid can be made smaller and smaller yielding approximate spatially continuous inference. We compare different approaches for producing grid level inference, based on mixture models and data augmentation, and compare their relative performances. We illustrate our framework with a new study on the spatio-temporal trends in respiratory disease risk in Glasgow, Scotland, and then move on to discuss some of the remaining challenges in this area of research.

A coupled model framework to detect nonlinear and delayed impacts of climate on dengue risk

Rachel Lowe
London School of Hygiene & Tropical Medicine

Extreme climatic events, such as floods and droughts, can impact the timing and intensity of epidemics of mosquito-borne diseases, including dengue and Zika viruses. Mosquitoes thrive in warm and humid conditions with rainfall increasing the number of outdoor breeding sites. However, drought conditions can also promote breeding in areas without a permanent water supply due to an increase in improvised water reservoirs. In this study, we couple distributed lag non-linear models (DLNM) with a spatio-temporal hierarchal mixed model framework. We fit the model to dengue case data across Brazil for the period 2001-2012 to understand the exposure-lag response association between dengue relative risk and temperature, extreme rainfall and drought events. Spatio-temporal random effects are included to account for unobserved confounding factors and spatial and temporal dependency structures. The model parameters are estimated in a Bayesian framework using integrated nested Laplace approximations in R (R-INLA). The model is fitted in cross-validation mode to produce out-of-sample probabilistic predictions of exceeding pre-defined epidemic thresholds and evaluated against a baseline model, representative of standard dengue surveillance practice. The advantage of our approach is the ability to capture cumulative and combined effects of anomalous climate conditions in the months leading up to a dengue epidemic. The coupling of DLNM with R-INLA is likely to prove an invaluable tool for assessing the impact of climate on health, to provide early warnings of the likelihood of epidemics and allow the effective communication of the associated uncertainty in probabilistic health risk forecasts to decision-makers.

Estimating under-five mortality and the coverage of health service provision in Kenya: A subnational spatio-temporal analysis, 1965-2015

Peter M Macharia[1], Emanuele Giorgi[2], Pamela N. Thuranira[1], Noel K. Joseph[1], Benn Sartorius[3], Robert W Snow[1,4], Emelda A Okiro[1]
[1]KEMRI Wellcome Trust Research programme, Nairobi
[2]Lancaster University
[2]University of KwaZulu-Natal
[4]University of Oxford

Background
Despite significant decline in under five mortality (U5M) over the last 3 decades, Kenya did not achieve Millennium Development Goal 4 (MDG 4) by 2015. We analyzed birth histories data to better understand spatio-temporal trends, inequalities of U5M and progress towards achieving reductions targets at decentralized and subnational decision-making units (counties) in Kenya between 1965 and 2015.
Methods
Ten household surveys and three censuses with birth histories undertaken between 1989 and 2014 were assembled and demographic methods applied to estimate U5M per county by survey. To generate a single U5M estimate for year and county, a Bayesian spatio-temporal Gaussian process regression was fitted accounting for variation in sample size, surveys and demographic methods. Inequalities and the progress in meeting goals set to reduce U5M were evaluated subnationally.
Results
Nationally U5M reduced by 61·6% from 141·7 in 1965 to 54·5 in 2013. The declining U5M was uneven ranging between 19 and 80% across the counties with some years when rates increased. 25 counties achieved the World Summit for Children goals but no county achieved MDG 4. Despite a striking decline in the levels of inequality, there was 3·8 times difference between counties with high and low U5M rates by

2013.

Conclusions

Kenya has made huge progress in child survival, however, U5M remains high and heterogeneous across counties. Better use of the current resources through focused allocation is required to achieve further reductions, reduce inequalities and increase the likelihood of achieving SDG 3·2 on U5M by 2030.

Alcohol outlet density and alcohol-related hospital admissions in England

Ravi Maheswaran
University of Sheffield

Background and aims

Excessive alcohol consumption has a substantial impact on public health services. A key element determining alcohol availability is alcohol outlet density. This study investigated the relationship between on-trade and off-trade outlets and hospital admission rates in local neighbourhoods.

Methods

A national small-area level ecological study design was used, incorporating all 32482 Lower Layer Super Output census Areas (LSOAs) in England (42 million people aged 15+ years). Densities for six outlet categories (outlets within a 1-km radius of residential postcode centroids, averaged for all postcodes within each LSOA) were calculated. Main outcome measures were admissions due to acute or chronic conditions wholly or partially attributable to alcohol consumption from 2002/03 to 2013/14.

Results

There were 1,007,137 admissions wholly, and 2,153,874 admissions partially, attributable to alcohol over 12 years. After adjustment for confounding, higher densities of on-trade outlets (pubs, bars and nightclubs; restaurants licenced to sell alcohol; other on-trade outlets) and convenience stores were associated with higher admission rate ratios for acute and chronic wholly attributable conditions. For acute wholly attributable conditions, admission rate ratios were 13% (95%CI 11-15%), 9% (7-10%), 12% (10-14%) and 10% (9-12%) higher respectively, in the highest relative to the lowest density categories by quartile. For chronic wholly attributable conditions, rate ratios were 22% (21-24%), 9% (7-11%), 19% (17-21%) and 7% (6-9%) higher respectively. Supermarket density was associated with modestly higher acute and chronic admissions but other off-trade outlet density was only associated with higher admissions for chronic wholly attributable conditions. For partially attributable conditions, there were no strong patterns of association with outlet densities.

Conclusions

Higher densities of several outlet categories were associated with higher hospital admission rates for wholly attributable conditions. Local licensing decisions may have an impact on the health of the local population and the burden on local hospitals.

Environmental Health: A LiDAR Assessment of Urban Neighbourhoods

MD Mahsin and Rob Deardon
University of Calgary

Modelling of infectious diseases has been increasingly used to evaluate the potential impact of different control measures and to guide public health policy decisions. In recent years, individual-level models (ILMs) have been effectively used to model infectious disease transmission. These models are well developed but assume the probability of disease transmission between two individuals depends only on their spatial (or network-based) separation. In this study, we extend ILMs to geographically-dependent ILMs (GD-ILMs) that allow the evaluation of the effect of spatially varying risk factors (e.g., education, social deprivation), environmental factors (such as temperature, air quality, rainfall, and humidity), as well as unobserved spatial structure, upon the transmission of infectious disease. We consider a conditional autoregressive (CAR) model to capture the effects of unobserved spatially structured latent covariates or measurement error. We show how GD-ILMs can be fitted to data on both simulation and Alberta influenza outbreaks epidemic within a Bayesian statistical framework using Markov chain Monte Carlo (MCMC) methods.

A geospatial analysis of childhood immunisation and associated factors: a cross-sectional study in New Zealand

Lukas Marek, Matthew Hobbs, Melanie Tomintz, Simon Kingham and John McCarthy
University of Canterbury

Immunisation is a simple, safe and effective way of protecting children and adults against harmful diseases before they are exposed to them. Childhood immunisation rates in New Zealand have increased steadily since the introduction of the National Immunisation Register (NIR) in 2005. However, there is a variation in immunisation rates at a sub-national level. Current indications suggest that the immunisation coverage of children is actually declining in some parts of the country. This study combined data from the NIR and other authoritative data sources, to identify well and insufficiently immunised areas. The study also explores associations (and their spatial variations) between immunisation coverage and a wider range of factors, including area-level socioeconomic deprivation, demographic characteristics and access to health-related services. The study shows that:

- significant differences in immunisation coverage exist, both overall and spatially, and that these variations remain when socioeconomic deprivation, demographic variables, health service accessibility, and urban/rural classifications are controlled for;
- the clusters of areas with low rates of immunisation appeared mainly in parts of the more densely populated areas of Auckland, Hamilton, Tauranga and Christchurch as well as in the central North Island regardless of the milestone age;

- higher immunisation rates are related to less socioeconomically deprived areas;
- the rates of ambulatory sensitive hospital (ASH) admissions of children (0–4 years) are lower in areas with higher immunisation rates;
- the accessibility to health-related services does not have a substantial impact on immunisation rates; as demonstrated by rural areas having higher immunisation rates than urban areas.

Field Sampling strategy for spatial distribution modelling

`Cedric Marsboom`
`Avia-GIS`

Field sampling is an expensive part of a research project. The correct planning of the sampling strategy, i.e. defining the number of samples and their location, is therefore important to both reduce cost and personnel effort. A successful sampling strategy depends on several parameters: type of study (transversal vs. longitudinal), type of data you want to collect (presence/absence vs. abundance), characteristics of the study area etc.

Whilst increasing the sample size, e.g. the number of locations to sample for a P/A study, may improve model outputs, this will substantially increase the amount of man hours and costs. Furthermore, above a certain threshold this will only provide limited gains in modelling accuracy.

In this presentation we focus on longitudinal presence/ absence studies. Using a virtual vector with a known spatial distribution we investigated the correlation between sample size and its spatial distribution on the one hand and spatial species distribution modelling on the other hand. We looked for the best trade-off between sample size and the stability of spatial model outputs and compared a series of accuracy statistics to determine the most cost-efficient sampling approach.

The relationship between children's objectively measured time spent outdoors, contact with nature, and emotional and behavioural resilience

`Paul McCrorie`
`University of Glasgow`

It is commonly assumed that children's interactions with nature have a positive impact on physical and mental health. Yet little evidence exists combining objective measures of environmental exposure, mobility, and nature. Here, we describe cross-sectional relationships between children's objectively measured time spent outdoors, contact with the natural environment, and emotional and behavioural resilience using Goodman's Strength and Difficulties Questionnaire (SDQ).

Studying Physical Activity in Children's Environments across Scotland (SPACES) was a large-scale project that collected 8-days of high-resolution (every 10 seconds) time-space-activity data (using accelerometry and GPS) from over 1200 children aged 10-11 years old during the period between May 2015 and May 2016. Children were part of a nationally representative birth cohort from the Growing Up in Scotland (GUS) study.

We linked demographic and survey responses from GUS, including total difficulties and sub-domain scores of the SDQ. Using the timestamped GPS data and Geographic Information Systems (GIS) software, we classified each 10-second as indoors or outdoors (using a GPS recorded quality variable), and the land use in which the GPS point fell.

Following data cleaning, multilevel analyses were conducted across the full sample (n=692), tested for moderating effects of sex, and any mediating impact of land use exposure between time spent outdoors and SDQ total/sub-domain scores. Models controlled for a number of individual, household, and neighbourhood factors, and were adjusted for the complex survey design of GUS. Results will be presented and discussed in relation to potential intervention development/evaluation and their implications for policy and planning.

Forecasting based on surveillance data

Sebastian Meyer
Friedrich-Alexander-Universität Erlangen-Nürnberg

Forecasting the future course of epidemics has always been one of the main goals of epidemic modelling. The Epidemic Prediction Initiative launched by the Centers for Disease Control and Prevention in the USA has triggered several seasonal influenza forecasting competitions. We review methods to evaluate the predictive performance of different statistical models for infectious disease surveillance time series. A case study demonstrates how to apply proper scoring rules to assess calibration and sharpness of predictive distributions with respect to various forecasting targets. The best model to forecast short-term incidence is not necessarily the best choice for long-term forecasts such as the peak week.

A Bayesian factor model for spatial panel data

Maura Mezzetti[1] and Samantha Leorato[2]
[1]Universitá Tor Vergat
[2]University of Milan

We propose a hierarchical Bayesian factor model for multidimensional spatio-temporal data that is able to take into account heterogeneities in latent factors (both in the time and space dimension) and allows further for a general form of autocorrelation in the observables, in all dimensions: temporal, spatial and between variables. This is obtained through a separable covariance structure, combined with an autoregressive structure for the latent vectors at every location. The proposed model is an extension of a model proposed by Mezzetti (SMA, 2012, pp.49–74) using the results of separable covariance matrix for the spatial panel data as in Leorato (Bayesian Analysis, 2016, 1-35). A Gibbs sampling algorithm is implemented to sample from the posterior distributions. The great flexibility of the Bayesian approach allows the inclusion of prior opinions about adjacent regions having highly correlated observable and latent variables. Data collected refer to the period 2006 - 2014 and to 27 European countries and consist in 17 variables indicating socio economic status, health indicators and sustainability are used to show the ability of our model in reduce the dimensionality of data. The ability to identify spatial patterns of disease risk that persist or evolve systematically over time provides more convincing evidence of true variations than a single crosssectional analysis. If multiple diseases are studied simultaneously, finding similar geographical or temporal trends of risk may further strengthen the evidence for common sources of influence that reflect underlying shared risk factors.

To what extent do Green and Blue Spaces improve our mental health and wellbeing?

`Amy Mizen`
`Swansea University`

Introduction
Spending time in green-blue spaces (GBS) has beneficial impacts on mental health. However, the evidence base is limited with respect to longitudinal studies. The main aim of this longitudinal, population-wide, natural experiment, is to quantify the impact of exposure to GBS on wellbeing and common mental health disorders, for a national population.

Methods
We developed a GBS typology in partnership with planners and policy makers. Using this typology, we built a national, longitudinal GBS dataset using local government audits and satellite data for 1.4 million homes in Wales, UK. We will create quarterly household access to GBS for eleven years (2008-2018). We will link GBS access scores to individual level mental health for 1.7 million people with primary care data and survey data (n = 12,000) on wellbeing.

Results
We created a national, longitudinal dataset of GBS and a cross-sectional dataset of access to GBS for 2018. Access to GBS varied by socio-economic status, urban/rural classification and type of GBS. We will expand this analysis to create quarterly household level GBS access for 2008-2018. We will explore how house moves, and visits to GBS change the association between GBS density and outcomes.

Discussion
This novel study fills the gap in the evidence base around environmental planning policy to shape living environments to benefit health. It will inform the planning and management of GBS in urban and rural environments and contribute to international work on impacts of the built environment on mental health and wellbeing.

Spatial Smoothing in Stan: insights from implementations of the ICAR prior and the BYM2 model

Mitzi Morris
Columbia University

In this talk, I'll introduce the Stan probabilistic programming language and its approach to coding the Besag-York-Mollié intrinsic conditional autoregressive (ICAR) model of additive heterogeneous and spatial random effects.

The combination of Stan's domain specific programming language and the Stan math library's rich set of well-tested and efficient implementations of probability distributions and matrix operations provides a middle ground between writing a custom model from scratch in a general-purpose programming language and using a fixed, pre-packaged implementation from an R package or Python library. This middle ground allows us to explore the consequences of different choices of priors, parameterizations, constraints, and similar implementation details.

To illustrate, I present different ways to implement constraints on the ICAR prior: either as a hard sum-to-zero constraint or using soft-centering prior. The soft-centering approach accelerates inference an order of magnitude over the usual sum-to-zero encoding. As a further example of Stan's expressiveness, I outline the implementation of the BYM2 model in Stan and extend it to Sterrantino et al's proposal for disconnected subcomponents.

Antibiotic Prescribing Patterns in General Medical Practices in England: Does Area matter?

Anna Molter
University of Manchester

Antimicrobial resistance is an important public health concern. In England, most antibiotics are prescribed in primary care; therefore, understanding prescribing patterns in General Medical (GP) practices is vital. The primary aim of this study was to spatially analyse patterns of antibiotic prescribing rates in GP practices in England. The secondary aim of this study was to examine the association of potential clusters of high or low prescribing rates with area level socio-economic deprivation using the Index of Multiple Deprivation (IMD).

This study used open source data from NHS Digital, the Office for National Statistics and Ordnance Survey. The spatial patterns of the annual mean prescribing rate of oral and intravenous antibacterial drugs per GP practice were analysed using the Getis-Ord Gi* statistic in ArcDesktop 10.4.1. GP practices belonging to clusters of high and low prescribers were extracted and their patient catchment IMD was compared in a spatial regression analysis using the spdep package (v0.7-4) in R v3.4.3. The pattern analysis identified a number of high and low clusters of antibiotic prescribing, with high clusters predominantly in the North of England. Spatial regression showed that patient catchments of high clusters were significantly more deprived than patient catchments of low prescribing clusters, especially in the domains of income, employment, education and health.

This study suggests the presence of area level drivers resulting in clusters of high and low prescribing. Consequently, area level strategies may be needed for antimicrobial stewardship rather than national level strategies.

A multivariate space-time model for trend examination and prediction of acute respiratory infection

Gary Napier
University of Glasgow

Surveillance systems for diseases monitor the rate of events on a daily, weekly, or monthly basis depending on the frequency of the event. For example monitoring for influenza in Europe occurs on a weekly basis though in Scotland there is daily monitoring of the death rate. Traditional systems, currently in use, tend to monitor the changes in the rates at a national level, though often with investigations at subnational level, such as health board or region. Here we will be examining the rates of acute respiratory infection (ARI) in Scotland across different age groups. We develop a Bayesian space-time regression model which permits modelling of a multivariate time series of counts within each spatial region. The multivariate nature could be rates of the same disease in different age/gender groups or different disease rates such as consultations for acute respiratory illness and consultations for gastro intestinal problems. The advantage of this approach is that we can incorporate both spatial and temporal correlation into the predictions of future events which are used for surveillance.

Distance sampling for epidemiology: estimating under-reporting of cases from clinic data

Luca Nelli
University of Glasgow

Distance sampling methods are widely used in ecology to estimate and map the abundance of animal or plant populations from spatial survey data. The key underlying concept in distance sampling is the detection function, a statistical description of the probability of detecting an organism of interest which can be expressed as a function of its distance from an observer and other covariates. In public health applications, data on disease incidence recorded at health clinics suffer from under-reporting bias of a similar nature, thus there is potential to adapt and apply methodology from ecology for epidemiological study. Here, we propose a modified framework of distance sampling in which some fundamental assumptions are relaxed so it can be applied to epidemiological data. We illustrate the approach by estimating variations in the probability of reporting malaria infection at clinics in a rural area of Burkina Faso, based on distances from villages according to the current road network. Finally we provide an interactive mapping tools (R Shiny app) that can be used by non-specialists to interrogate model outputs and visualize how the overall probability of under-reporting and the catchment area is influenced by changing the spatial distribution of health centres.

Recent rises buck long-term reductions in subnational Infant Mortality Rates: UK 1981 to 2017

Paul Norman
University of Leeds

This work looks at the intersection of long-term spatial trends of area deprivation and trajectories of infant mortality; a key outcome at the start of the life-course. Despite substantial reductions in Infant Mortality (deaths to infants aged less than one year) over the last few decades, improvements at national level have slowed compared to other countries and a rise in IMR has recently been observed. This study utilises UK coverage Vital Statistics data for 1981 to 2017 to determine subnational trends in infant mortality.

After reporting national trends for the UK's countries, a time-series of subnational IMRs are presented, classified into trajectory groups to identify those areas who are ahead of the game and those who are getting left behind. These IMR trend groups are cross-tabulated with places classified by their deprivation trajectories; whether persistently deprived, advantaged, improving or not. A rise in infant mortality is evident during 2013-17 for 36% of local authorities so this is explored to highlight the geography of where IMRs have risen. Disaggregating infant mortality by neonatal (less than 28 days) and postneonatal (28 days to 1 year) shows that the recent rise is largely accounted for by neonatal deaths.

Life course neighbourhood deprivation effects on body mass index: quantifying the importance of selective migration.

Paul Norman[1] and Emily Murray[2]
[1]University of Leeds
[2]University College London

Neighbourhood effects research is plagued by the inability to disentangle effects of neighbourhoods from selection sorting of people into neighbourhoods over time. We used data from two British Birth Cohorts, the 1958 (ages 16, 23, 33, 42 and 55) and the 1970 (ages 16, 24, 34 and 42), and structural equation modelling, to investigate life course relationships between body mass index [BMI] and area deprivation (addresses at each linked to the closest census 1971-2011 Townsend score [TOWN] re-calculated to reflect 2011 lower super output level (LSOA) boundaries). Initially, model fit was compared between three models: (1) area deprivation only, (2) health selection only and (3) both combined. Following, in the best fitting model relationships were tested for effect modification by residential mobility by inclusion of interaction terms. Findings showed that model fit statistics were contradictory, so the most complex model – with both area deprivation and health effects included – was chosen for further analysis. For both cohorts, both BMI and area deprivation tracked across the life course. In both cohorts, relationships occurred between area deprivation and BMI at the next interval at all age intervals except in NCDS, TOWN23->BMI33. In contrast, health selection paths only occurred at three intervals (NCDS: BMI23->TOWN33 and BMI33->TOWN42; BCS: BMI34->TOWN42). Moving between study intervals was related to BMI and TOWN at some ages, but did not modify any area deprivation effect or BMI health selection paths. In conclusion, while selective sorting by BMI does occur in later adulthood, it does not explain adolescent and adulthood associations between area deprivation and BMI.

Relationships between individual and urban characteristics and mobility for 122 individuals living in Antwerp, Barcelona and London

Jonathan Olsen
University of Glasgow

It remains standard practice for researchers to assume someone's immediate neighbourhood of residence is a good proxy for their environmental exposures, the amenities and facilities they use and where best to target them for intervention. However, the advent of Global Positioning Systems (GPS) devices which track people has revealed that most individuals have a far more complex space-time geography and that the composition and configuration of the wider environment affects where and how people move within it. We explored relationships between individual characteristics (age, sex, socio-economic status and mobility restrictions), urban characteristics and mobility for 122 individuals living in Antwerp, Barcelona and London from the Physical Activity through Sustainable Transport Approaches (PASTA) study (www.pastaproject.eu). Individuals wore GPS devises for one week during three separate occasions during a 12 month period. Firstly we will describe methodological challenges in using GPS datasets that yield large amounts of precise location data using the appropriate spatial scale. Secondly, we will provide initial findings describing variation of individual mobility and environmental exposures by city, weekday/end and individual characteristics.

The influence of space and social mix on patterns of urban green space usage. An Agent-based model

Stefano Picascia
University of Glasgow

Evidence from surveys and other sources shows a complex patterning in the usage of urban green space. Age and gender differences, plus the ubiquitous class gradient, don't seem to account for the whole variability in the data. Moreover, wide differences exist between cities, with residents of certain urban areas using their local green spaces orders of magnitude more than others. We propose a simple spatially explicit agent based model to investigate the extent to which social influence and social cohesion, in their interaction with specific urban forms, might affect the emergence of a shared culture of utilizing (or not utilizing) public parks.

Using propensity score to adjust for residual confounding in area-referenced environmental health studies

Monica Pirani
Imperial College London

Study designs where data has been aggregated by geographical areas are popular in environmental epidemiology. These studies are commonly based on administrative databases and, providing a complete spatial coverage, are particularly appealing to make inference on the entire population. However, the resulting estimates are often biased and difficult to interpret due to unmeasured confounders, which typically are not available from routinely collected data. We propose a framework to improve inference drawn from area-referenced environmental health studies exploiting information derived from individual-level survey data.

In particular, our proposed approach (i) summarises available individual-level confounders into an area-level scalar variable, which we call ecological propensity score (EPS), (ii) implements a hierarchical structured approach to impute the values of the EPS in the areas where they are missing under different missingness assumptions, and (iii) includes the estimated and imputed EPS into the ecological regression model that links the risk factors of interest to the health outcome.

This delivers area level risk estimates which are less biased, allowing a fuller adjustment for confounding than traditional areal studies. By using simulated examples, we show that such a way of building and using EPS is a promising method to reduce the bias intrinsic in area-referenced studies due to unmeasured confounders, and we demonstrate its benefits for both binary and continuous exposure. The approach is applied to case studies in environmental epidemiology, investigating the effect of ambient air pollution on health outcomes in England (UK).

Alleviating spatial confounding for areal data problems by displacing the geographical centroids

Marcos Oliveira Prates
Universidade Federal de Minas Gerais

Spatial confounding between the spatial random effects and fixed effects covariates has been recently discovered and showed that it may bring misleading interpretation to the model results. Techniques to alleviate this problem are based on decomposing the spatial random effect and fitting a restricted spatial regression. In this paper, we propose a different approach: a transformation of the geographic space to ensure that the unobserved spatial random effect added to the regression is orthogonal to the fixed effects covariates. Our approach, named SPOCK, has the additional benefit of providing a fast and simple computational method to estimate the parameters. Also, it does not constrain the distribution class assumed for the spatial error term. A simulation study and real data analyses are presented to better understand the advantages of the new method in comparison with the existing ones.

Illness and the lifecourse: does the relationship vary by cohort?

Lucy Prior
University of Bristol

Understanding trends of health states over time is critical for furthering our understanding of the development of illness and health inequalities. Analyses of health over time must consider the impact of ageing as well as the potential for cohort differences in health trajectories. The British Household Panel Survey (BHPS) and Understanding Society longitudinal studies are employed to assess self-rated health and mental health trends over a 26-year period. This analysis uses cross-classified multilevel models in an exploratory, non-parametric approach to evaluate age and cohort effects net of each other. In line with previous research, we find self-rated health tends to worsen over the lifecourse, following a peak in middle-age. However, younger cohorts show improved self-rated health compared with older generations at similar age-points. Mental health evidences a trend of initial decline as people age which reverses and then exhibits improvement in late-middle-age, with less defined cohort trends. We additionally report on interactions between age and cohort trajectories. In particular, we identify young adults born in 1990s as presenting better self-rated health than their peers from older cohorts when assessed at the same age, with further evidence to support an additional negative burden of mental health for the youngest cohorts, although this is much weaker. Ultimately, we suggest that researchers would benefit from wider use of this exploratory modelling strategy when evaluating underlying health trends and more research is now needed to explore potential explanations of these baseline trajectories.

Real Time Decision Making for Infectious Disease Outbreaks

Will Probert[1], Mike Tildesley[2]
[1]University of Oxford
[2]University of Warwick

In the event of outbreaks of infectious diseases, mathematical models can be used to inform decision makers regarding the likely spread of disease and the impact of control strategies. However, in the early stages of novel outbreaks, there can often be significant uncertainty regarding the spatiotemporal spread of disease and the likely impact of any intervention policy. However, policy makers do not often have the luxury to wait for any uncertainty to resolve before introducing an intervention, so it is crucial that models are developed to take account of the most up to date information available. We analyse historical outbreaks of foot-and-mouth disease and avian influenza and investigate the predictability of infectious disease models in the early stages of these epidemics to determine how any predictions change as more data are accrued. Our results indicate that substantial epidemiological uncertainty at epidemic onset can lead to misleading forecasts of the impact of any intervention policy. However, robust predictions can be obtained after the first two to three weeks owing to a resolution of uncertainty during this period. We conclude that real time information is vital to ensure that policy makers can select the most appropriate intervention policy to minimise the impact of any ongoing epidemic.

Mapping Trypansoma cruzi infection prevalence

`Andre Python`
`University of Oxford`

American trypanosomiasis, or Chagas disease, is a leading cause of heart disease in Latin America. This neglected tropical disease is caused by infection with the Trypanosoma cruzi parasite, which is primarily transmitted to humans by triatomine vectors, blood-feeding insects that feed on mammals, marsupials, birds and reptiles. This project aims to map at pixel-level the overall prevalence of T. cruzi infections in vectors in order to target vector control within the endemic zone. Pixel-level species-specific intensity maps of vector species occurrence are produced using Poisson point process models based on presence-only data. These are then used in a data augmentation step that combines point-level and polygon-level infection prevalence records. Spatial, block-wise cross validation techniques are employed for model evaluation. The final result is a predictive map of mean infection prevalence in vectors with associated uncertainty.

Trends in Tract-Level Obesity Rates in Philadelphia by Race/Ethnicity, Space, and Time

`Harrison Quick`
`Drexel University`

The growing recognition of often substantial neighborhood variation in health within cities has motivated greater demand for reliable data on small area variations in health outcomes. The goal of this work is to investigate both racial and geographic disparities in obesity rates in Philadelphia Census tracts over the period 2000–2015. Our data consist of self-reported survey responses of whites, blacks, and Hispanics from the Southeastern Pennsylvania Household Health Survey. To analyze these data – and to obtain more reliable rate estimates – we apply a multivariate spatiotemporal Bayesian model, simultaneously accounting for spatial-, temporal-, and between-race/ethnicity dependence structures. This approach yields estimates of the obesity rate by age, race/ethnicity, sex, and poverty status for each census tract at all time-points in our study period. While the data suggest that obesity rates have increased at the city-level for men and women of all three race/ethnicities, the magnitude and geographic distribution of these increases differ substantially by race/ethnicity and sex. The method can be flexibly used to describe and visualize spatial heterogeneities in levels, trends, and in disparities. This is useful for targeting, surveillance, and to motivate etiologic research.

Estimating the impact of air pollution in Scotland and potential benefits of pollutant concentration reductions

`Colin Ramsay`
`Health Protection Scotland`

Background
Air pollution levels in Scotland are generally low except for elevated pollutant concentrations in city centres, especially NOx/ NO2. Low Emission Zones (LEZs) using EU vehicle emission standards to restrict access to affected areas by the most polluting vehicles have been adopted as an intervention. Estimating the potential health benefits of an LEZ requires understanding of the underlying relationship between pollutant concentrations and measures of health impact. Recent data on these relationships at local level in Scotland was lacking. This pilot study set out to examine these relationships as a possible basis for quantifying the health impact of an LEZ.
Methods
Data on cardio-respiratory morbidity and mortality, as well as total non-accidental disease mortality were reviewed for a two year period (2015 to 2016) for 1252 spatially defined Intermediate Zones with an average 4000 population and linked to air pollution data using modelled annual average pollutant concentrations for the relevant years for NO2, NOx, PM10 and PM 2.5. The impact of air pollution on health outcome risk was analysed using a spatial hierarchical regression model (Lee and Mitchell 2013) controlling for covariates derived from the Scottish Index of Multiple Deprivation (SIMD). Single pollutant, single outcome models were used to account for high pollutant collinearity providing 20 disease-pollutant combinations.
Results
Significant associations were found with respiratory hospital admissions for all pollutants and for respiratory mortality with PM2.5. Cardiovascular disease outcomes, nor total non-accidental mortality showed significant association with any pollutant. Resulting estimates showed LEZs may have relatively modest health benefits.

Modelling spatial distribution of ticks in Scotland

Rita Ribeiro
Scotland's Rural College

Ixodes ricinus is the most abundant and widespread tick species in Western Europe and is the vector of several pathogens, including the complex of Borrelia bacteria that causes Lyme disease. Although tick distribution maps are valuable for public health decision-making, modelling I. ricinus distribution is challenging, due to the complex tick ecology and the difficulty of getting robust tick data at a national scale. In Scotland maps of I. ricinus distribution at an appropriate scale for national and regional decision-making are currently limited. Here we show the results of predictive models and maps for I. ricinus abundance and presence-absence, which were developed using the most extensive datasets of this tick species that are available for Scotland, comprising data from public submissions and from scientific quantitative surveys. Although the models fit the data well and the significant environmental predictors are consistent with the published literature, the predictions over mainland Scotland are associated with a large amount of uncertainty, particularly for the model fitted with tick presence data from public submissions. The main limitation of using tick count data from scientific surveys was the uneven geographical distribution from which the tick count data were collected, so that predictions were less accurate when extrapolating to areas with poor data coverage. We used the findings to explore the robustness of predictions, focusing on errors and uncertainty resulting from data quality. This illustrates the challenges of modelling vector distribution that can arise due to incomplete data, and we make suggestions to guide future studies.

Public Health Applications of a Bayesian space–time model for clustering areal units based on their disease trends

Chris Robertson[1], Duncan Lee[2] and Gary Napier[2]
[1]Strathclyde University
[2]University of Glasgow

There is public health concern about the stalling life expectancy in Scotland, and other UK nations, over the last 5 years [1] and some authors have attributed this to an effect of austerity measures which have been in place since about 2010. It is also well established that there are spatially aggregated mortality differences over Scotland with lower mortality rates in the more affluent parts of Scotland [2]. This analysis will investigate the hypotheses that there are clusters of areas, which have seen no change in the decline of mortality, and other areas where morality rates have been increasing since 2010. The analysis is carried out using a recently developed a clustering paradigm to identify groups of areal units exhibiting similar temporal disease trends. [3] The detection of clusters of areas abased upon their trends is based upon a novel Bayesian hierarchical mixture model. Inference is based on a Metropolis-coupled Markov chain Monte Carlo ((MC)3) algorithm. This model is included within the CARBayesST package in R [4]. The results of this analysis show the implementation of sophisticated space time models within a public health setting and discuss how the outputs of these models can be used to inform policy.

1. https://www.biorxiv.org/content/10.1101/542472v1
2. https://www.scotpho.org.uk/comparative-health/health-inequalities/policy-context/
3. Napier G, Lee D, Robertson C, Lawson A. A Bayesian space–time model for clustering areal units based on their disease trends. Biostatistics. 2018
4. https://cran.r-project.org/web/packages/CARBayesST/vignettes/CARBayesST.pdf

When theory and practice collide: A study of ground ambulance bases in Southern Norway

Jo Røislien
University of Stavanger

Emergency medical services (EMS) provide quick responses in emergency situations. When a recent mathematical study of air ambulance base locations in Norway indicated that by radically changing the current base locations coverage could be increased, even with significantly fewer bases, it was received with mixed feelings. The gap between mathematical theory and clinical practice became a hot topic.

The Vestfold region in Norway is situated southwest of the capital Oslo, covering 2224 km2 with approx. 250,000 inhabitants. The EMS in the region has five fixed ambulance bases operating 13 ambulances during the day and eight during nighttime. Using a modified Maximum Expected Covering Location Model (MEXCLP) the optimal location of bases, and corresponding distribution of available ambulances, was estimated. Numerous models were fitted, for varying numbers of ambulances and ambulance utilizations. Mathematically optimal results matched the location of existing bases surprisingly well. In a greenfield scenario the mathematical optimal solution spread out the available ambulances, but still in tight clusters around the existing base structure.

The location of the northernmost base was troublesome. Statistics Norway classifies the city of Holmestrand as rural, and the mathematical models placed the northernmost base far from the city. This is in stark contrast to how Holmestrand is viewed by the region's inhabitants. Modifying Statistics Norway's urban/rural classification the mathematical models approached the existing base structure. This demonstrates the importance of correct classification of demand points, and national guidelines might not be optimal. The study highlights the importance of close collaboration between mathematicians and practitioners.

Distributed lagged modelling for spatio-temporal areal data with excessive zeros

Chawarat Rotejanaprasert
Mahidol University

In much of the Greater Mekong Sub-region, malaria is now confined to patches in very specific areas. The malaria landscape is quite heterogeneous, with many of these patches associated with environmental factors. An aim of malaria surveillance program is to identify focal areas vulnerable to disease transmission. However, the possible effect at different lag periods between meteorological variables and clinical malaria has not been well studied in the region. Thus in this study we developed distributed lagged modelling accounting for the lagged effect and excessive zero cases to facilitate in malaria elimination planning. A case study was provided with weekly data of environmental factors and clinical malaria cases in Thailand.

Utilising an 'adaptive raster' approach to reduce uncertainty in sub-national c-section estimates in Tanzania, 2015

Corrine Warren Ruktanonchai
University of Southampton

Spatial inequalities in access to maternal and newborn health (MNH) services persist both within and between countries, marginalizing some of the world's most vulnerable populations. Identifying and targeting these vulnerable populations is a top Sustainable Development Goal priority, requiring model-based predictions of

MNH services at small spatial scales. Global estimates have historically relied on aggregate household-level surveys, such as the Demographic and Health Surveys, which often employ complex hierarchical sampling design, resulting in model uncertainty that requires effective communication to policy makers and non-academic decision makers. This uncertainty propagates heterogeneously across space and impacts model precision, influenced by the availability, spatial resolution, and amount of geo-spatial covariates.

Towards this, we employ a novel ‘adaptive raster’ approach allowing grid cells to vary non-uniformly across space, to better communicate and account for the associated uncertainty inherent to geo-statistical models. We assemble a spatial database of environmental and socioeconomic covariates linked with household surveys, including poverty indices, accessibility to health facilities, urbanicity, births, population, and travel time to the nearest city. We employ Bayesian hierarchical models to explore how uncertainty propagates at multiple spatial resolutions, and how average spatial resolution changes for modelled estimates using an adaptive raster approach. Lastly, we present a non-uniform modelled c-section surface with associated uncertainty, reflecting greater programmatic relevance in reaching global maternal health targets. These results have policy relevance in outlining where access to maternal health services is inequitably distributed across space, and methodological implications for datasets addressing rare events or having high model uncertainty.

Historical Air Pollution and Brain Health: Life Course Epidemiology Insights from the Lothian Birth Cohort 1936

Tom Russ
University of Edinburgh

Background: Air pollution has been consistently linked with dementia and cognitive decline. However, it is unclear whether risk is accumulated through long-term exposure or whether there are sensitive/critical periods. One barrier to clarifying this is the dearth of historical air pollution data. Methods: Using the EMEP4UK chemical transport model[1] we modelled historical particulate matter (PM) for the years 1935, 1950, 1970, 1980, and 1990 and combined these with contemporary modelled data from 2000 onwards to estimate life course exposure in 467 participants in the Lothian Birth Cohort 1936[2] with lifetime residential history recorded[3]. Linear regression models using cognitive change from [A] 11-70 years and [B] 70-76 years were constructed adjusting for the rich lifestyle, medical, and socioeconomic data collected in the LBC1936. Thus, we could explore when exposure to air pollution was most important in relation to brain health. Results: Preliminary results suggest that, for men, an accumulation of risk model fits the data best for both outcomes with some evidence for exposure in the fifth decade being most strongly associated with cognitive change from 70-76. In women, the accumulation model also fitted best with cognitive change from 11-70 but exposure two decades later was implicated; there was a suggestion that early life exposure may be more important for later life cognitive change in women. Conclusions: The life course paradigm is essential in understanding dementia and ours is the first study to examine life course air pollution exposure in relation to brain health.

[1] doi:10.5194/acp-14-8435-2014
[2] doi:10.1093/ije /dyr197
[3] doi:10.1016/j.socscimed.2017.10.038.

Grid square level inference in Bayesian disease mapping

Kamol Sanittham, Craig Anderson and Duncan Lee
University of Glasgow

Disease mapping focuses on estimating spatial variation in disease risk across areal units. However, the choice of the areal unit is arbitrary and therefore when the units are changed, the data also change and consequently change the results. This study outlines a methodology to transform the areal unit data into pseudo-continuous grid square level data to overcome this problem. The aim of this research is to estimate the spatial pattern in disease risk at the grid square level based on areal unit data. There are two different approaches proposed to estimate disease risks at the grid level, utilising multiple imputation and data augmentation. Both approaches utilise conditional autoregressive models to capture spatial correlation, and inference is carried out via Markov chain Monte Carlo simulation. The efficiency of the methodology is confirmed by a simulation study, and is then illustrated by respiratory disease hospitalisation data in Glasgow, Scotland between 2015 and 2016.

Disease risk estimation in small areas exploiting the underlying clustering structure of the data

Guzman Santafe
Public University of Navarre

Many statistical models have been developed during the last years to smooth risks in disease mapping borrowing information from nearby regions. However, most of these modelling approaches do not take possible local discontinuities into consideration. Local discontinuities are very usual in real data analysis, especially when dealing with small areal units, and they are generally over-smoothed in disease risk maps. In previous proposals, a two-stage approach was considered to obtain a clustering partition of the areal units (first stage) and to estimate spatial risks by fitting Bayesian hierarchical models including the cluster configuration structures. Nonetheless, these proposals adapt standard hierarchical clustering algorithms to the spatial (or spatio-temporal) areal data context obtaining a set of “n” (number of areas) spatially contiguous cluster structure candidates to be evaluated. The computational cost of these approaches when dealing with a large number of small areas may be prohibitive. In this talk, we propose a novel density-based clustering algorithm designed to exploit the special characteristics of the spatial problem. In contrast to previous algorithms, the proposed method is able to automatically detect the number of spatial clusters and therefore, a single cluster structure is obtained. This makes the algorithm suitable for analysing large spatial datasets. A simulation study has been conducted to evaluate the performance of this new spatial clustering algorithm in comparison with a previous proposal, obtaining competitive risk estimates at much better computational cost. The new methodology will be applied to analyse stomach cancer mortality data in 8000 municipalities in Spain.

Estimation of health co-benefits gained through the reduction of air pollution through the implementation of global climate mitigation commitments

Gavin Shaddick
University of Exeter

The Paris Agreement, an agreement within the United Nations Framework Convention on Climate Change (UNFCCC) dealing with greenhouse gas emissions (GHG) mitigation, adaptation, and finance starting in the year 2020. The Agreement aims to keep the increase in global average temperature to below 2°C above preindustrial levels and to pursue limiting the increase to 1.5°C, which would reduce the risks and impacts of climate change. Air pollution is one of the most important environmental risk factors to global health and here we investigate the potential health co-benefits of climate change mitigation strategies through reductions in ambient fine particulate matter air pollution (PM2.5) under different climate and socioeconomic scenarios. This process comprises of four steps:

1. The Global Change Assessment Model is used to estimate GHG and air pollutant emission corresponding to different climate and socioeconomic scenarios.
2. Projected emissions are used as inputs to the TM5-FASST air quality source-receptor model to estimate effects on global concentrations of PM2.5
3. The Data Integration model for Air Quality is used to calibrate concentrations against ground monitoring data, producing country-level distributions of population-weighted concentrations.
4. Burden of disease is calculated, using the outputs from DIMAQ and the Integrated Exposure Response function.

Significant reductions in the expected number of deaths from causes associated with exposure to PM2.5 were observed in 2050 under a scenario with a long-term temperature stabilization target (2oC) when compared to a reference scenario, equating to ca. 4% of the total deaths associated with air pollution.

Neighbourhoods, selection bias and the MARMoT approach to compare unbalanced groups

Margherita Silan, Giovanna Boccuzzo and Bruno Arpino
University of Padua

The study of neighbourhood effects on health condition has gained attention exponentially during last decades. This topic is even more important from a public health perspective when the focus is on health conditions of elderly. Indeed, old individuals are more susceptible to neighbourhood effects because they spend more time in their neighbourhood than young people. The aim of our work is to estimate neighbourhood effects on elderly health outcomes in Turin (a city located in north Italy) using data coming from the Turin Longitudinal Study, that contains linked individual data from censuses and administrative health data flows. Individuals are not randomly distributed among neighbourhoods, the residential area is chosen also on the basis of individual socio-economic characteristics. This causes a selection bias, that may confound comparisons between the distribution of health outcomes among neighbourhoods. In order to balance observable confounders and to make different populations living in different neighbourhoods comparable, we adopted a propensity score approach. The original methodological contribution of our work is the adaptation of propensity score techniques to a framework with many treatments (that are neighbourhoods in this application). Indeed, Turin may be divided according to three geographical partitions in 10, 23 or 94 neighbourhoods. We proposed a novel method that consists on a Matching on Poset based Average Rank for Multiple Treatments (MARMoT), which has revealed to be really useful to improve the covariates' balance between groups and pretty fast with respect to other approaches, that have revealed to be unpractical in applications with many treatments.

A multivariate hierarchical model for the use of polluting fuels for cooking

Oliver Stoner and Theo Economou
University of Exeter

Around 3.8 million deaths per year are attributed to household air pollution worldwide. Information on the proportion of people relying primarily on different polluting fuels for cooking, which acts as a proxy for pollution exposure, is available in the form of nationally-representative household surveys. However, the absence of a modelling framework for comprehensively estimating the use of individual fuels inhibits fuel-specific policy interventions. To address this, we develop a multivariate hierarchical model for data from the World Health Organization Household Energy Database, spanning the period 1990-2016. Based on Generalized-Dirichlet-Multinomial distributions, the model jointly estimates trends in the use of eight individual fuels, whilst addressing a number of challenges involved in modelling the data. These include: missing values arising from incomplete surveys; missing values in the number of survey respondents; and sampling bias in the proportion of urban and rural respondents. The model also includes regional structures to improve prediction in countries with limited data. We assess model fit using within-sample predictive analysis and conduct an out-of-sample prediction experiment to evaluate the model's forecasting performance. Overall, this work substantially contributes to expanding the evidence base for household air pollution, which is crucial for developing policy and planning interventions.

Spatio-temporal Geostatistical Modelling to jointly Predict Exposure and Disease

Giorgia Stoppa and Dolores Catelan
University of Florence

We present a joint Bayesian space-time geostatistical model for exposure assessment and disease risk estimation. The epidemiological study was a panel study on children respiratory health in the environmental high risk area of Valle del Mela (Sicily, IT). Concentration measurements on Sulphur dioxide (SO2) were collected on 12 weeks in 21 locations by passive dosimeters on the period November 2007 – April 2008. Baseline information on risk factors and outcomes were obtained at beginning of the study period (crossectional study on 2506 children) and weekly diaries on outcomes were collected during the study on a panel of a subset of 125 asthmatic children (Biggeri et al. 2014 WHO ISBN: 978 9 289 05005 0). The Bayesian hierarchical model consists into two layer. The exposure model specifies a space-time Bayesian geostatistical models to predict pollutant concentrations at children's residential addresses. The model is relevant for exposure assessment based on point referenced data (see Biggeri et al. IBC 2018; Vicedo-Cabrera 2013 GeoSpatialHealth). The disease model consists in a random effect longitudinal logistic regression model. The original feature of this modelling is the joint specification of a spatial long-term effect and a spatiotemporal short-term effect of the pollutant concentrations (see for example (Kloog et al. 2013 Epidemiology).

Spatial scan statistics can be dangerous

`Toshiro Tango`
`Center for Medical Statistics, Japan & Teikyo University`

Spatial scan statistics are widely used tools for detection of disease clusters. Especially, the circular spatial scan statistic (SaTScan) proposed by Kulldorff (1997) has been utilized in a wide variety of epidemiological studies and disease surveillance. The flexible spatial scan statistic (FleXScan) proposed by Tango and Takahashi (2005) has also been used. However, it does not seem to be well recognized that these spatial scan statistics tend to detect the most likely cluster, much larger than the true cluster, by swallowing neighbouring regions with non-elevated risk. In this presentation, to avoid detecting these undesirable clusters, we shall illustrate the use of FleXScan with a restricted likelihood ratio proposed by Tango (2008) with simulated data and real mortality data.

Data integration for high-resolution, continental-scale estimation of air pollution concentrations

`Matthew Thomas`
`University of Bath`

Air pollution constitutes the highest environmental risk factor in relation to heath. In order to provide the evidence required to perform health impact analyses, to inform policy and to develop potential mitigation strategies comprehensive information is required on the state of air pollution. Traditionally, information on air pollution comes from ground monitoring (GM) networks but these may not always be able to provide sufficient coverage and may need to be supplemented with information from other sources (e.g. chemical transport models). However, these other sources may only be available on grids and, as such, cannot capture micro-scale features that may be important in assessing the levels of pollution in areas of high population. Here we develop a model that allows calibration between data sources available at different levels of support, allowing the coefficients of calibration equations to vary over space and time. Set within a Bayesian hierarchical framework we address the computational issues that can arise when fitting varying coefficient models in larger scale problems, especially those using Markov Chain Monte Carlo, by perform approximate Bayesian inference based on integrated nested Laplace approximations. The model is used to produce high-resolution (1km x 1km) estimates of NO2 and PM2.5 across Western Europe for 2010-2016. Posterior predictive distributions are produced for each grid-cell which offer a wealth of information, including the ability to calculate exceedance probabilities and the calculation of country-level population-weighted concentrations. Although levels of air pollution are decreasing, there remain large populations that are exposed to levels that exceed the WHO Air Quality Guidelines.

Investigating the variation of antibiotic prescribing rates between NHS Scotland health-boards and GP practices through the use of maps, principal components and spatial analysis

Florence Tydeman
University of Strathclyde

The data used in these analyses were open-sourced GP antibiotic prescription data from 2016. The use of Principal Component Analysis (PCA) on the 13 antibiotic groups (defined by BNF code) indicated GP practices that followed similar prescribing trends. A Poisson glm of total antibiotic prescribing assessed associations with GP descriptive information. Separate spatial data frames were constructed for health-boards and GP-practice level data. Areal maps were created to visualise rates of antibiotic groups and point maps for individual GP rates. Tests for spatial association were conducted for total antibiotic prescribing rates using Monte Carlo Envelopes for the GP-practice point data and Moran's I at health-boards level. Results from these tests introduced potential to build conditional auto-regressive (CAR) models to account for spatial variation.

This analysis has shown that the variation in antibiotic prescribing rate between GP practices can be partly explained by practice demographics. Increased proportions of over 74, under 15 and most deprived quintile suggests higher prescribing, whereas increasing proportion of least deprived suggests lower prescribing rates. The PCA has shown that 25% of the total variation can be explained by a weighted average across all antibiotic drug groups hence total antibiotics is a reasonable representation of prescribing at GP-level. After adjusting for GP practice characteristics, there is no evidence of spatial association between GP practices. Moran's I test for spatial association suggested associations between health-boards for antibiotic prescribing. This can therefore be investigated further through the use of CAR models.

Mapping vaccination coverage rates at high resolution to explore the effects of delivery mechanisms

Chigozie Edson Utazi
University of Southampton

The success of vaccination programs depends largely on the mechanisms used in vaccine delivery. National immunization programs offer childhood vaccines through fixed and outreach services within the health system and often, additional vaccination activities are undertaken to boost coverage and fill gaps in population immunity. The effects of these different delivery methods on coverage rates have not been explored at fine spatial scales across multiple settings. Here, we map coverage at 1 x 1 km spatial resolution in five case-study low- and middle-income countries using Demographic and Health Survey data. We compare estimates of the coverage of the third dose of diphtheria-tetanus-pertussis-containing vaccine (DTP3), which in these countries is delivered through routine immunization (RI) with those of measles-containing vaccine (MCV) for which campaigns termed supplementary immunization activities (SIAs) are also undertaken. Additionally, we examine different aspects of health and immunization system performance by comparing the coverage of the three doses of DTP vaccine and the corresponding dropout rates. We find that SIAs have been effective in boosting MCV coverage in some places, but not in others, particularly where RI had been deficient as depicted by DTP coverage. Also, dropout rates between the three-dose DTP series were highest in places with poorer coverage with the first dose. Our findings highlight the need for health and RI system strengthening, and that the modelling approaches outlined here can help to guide geographical prioritization and strategy design.

Mental wellbeing of older urban residents: the MINDMAP project

Frank J. van Lenthe
Utrecht University

Urbanization and ageing have enormous implications for public mental health. Cities pose major challenges for older citizens, but also offer opportunities for the design of policies, clinical and public health interventions that promote mental health. The overall aim of the Horizon2020-funded MINDMAP project is to identify the opportunities offered by the urban environment for the promotion of mental wellbeing and cognitive function of older individuals in Europe.

A core element of MINDMAP is the development of a data platform of 10 international harmonized longitudinal urban cohorts of ageing covering 16 cities in Europe, Canada and the US, which will be linked to an innovative database of urban characteristics and policies. Pathways linking the urban environment to mental health and cognition via lifestyle behaviour, social influences and mobility patterns in the built environment will be investigated, thereby exploiting the variation across cities. Agent-based models will be used to simulate policies and prevention strategies on mental health in urban environments.

It is the aim of this presentation to (1) describe the design of the project, (2) discuss the methodological challenges linking urban characteristics of European cities to health outcome and (3) present some initial findings of the project.

Estimating the spread of tick-borne encephalitis virus in ticks and hosts

Melanie Walter, Janna R. Vogelgesang, Franz Rubel and Katharina Brugger
University of Veterinary Medicine Vienna

Each year, more than 2,000 human tick-borne encephalitis (TBE) cases are confirmed in Europe, a remarkable proportion outside official risk areas or know endemic areas. The population in non-risk areas increasingly ask, whether they should consider a vaccination against TBE. Available risk maps, like those of the Robert Koch-Institute in Germany, are based on registered human TBE cases. Alternatively, the risk of being bitten by a TBE virus infected tick (exposure-based approach) can be estimated with a species distribution model. Contrary to human TBE case-based risk assessment, the spatial risk is estimated based on virus detections in ticks and animal hosts. Therefore, a Random Forests model based on decision trees was implemented and applied to a dataset comprising more than 800 georeferenced TBE virus records in ticks and animal hosts. As environmental predictors temperature- and precipitation-dependent bioclimatic parameters of the Worldclim dataset as well as land use and coverage of the GlobCover dataset were applied. With the Random Forests model the potential TBE virus distribution in Central Europe is estimated. Beside in known endemic areas, high habitat suitability is also shown in northern Germany, the Netherlands, and Belgium. This novel map can support decision makers to identify risk areas for human TBE infections.

Exploring patterns of association between Experian's Mosaic classification and all-cause mortality in England: implications for further research

Welcome Wami, Oarabile Molaod, Ruth Dundas, Alastair Leyland and Vittal Katikireddi
University of Glasgow

Background
The Index of Multiple Deprivation (IMD) measures relative deprivation in England. It is widely used by central and local government to distribute funding, target resources or prioritise delivery of interventions to areas. Following the move of public health services to local government, there is considerable interest in using additional small area socioeconomic measures that may be amenable to change to improve local responsiveness. Our study explored relationships between Experian's Mosaic geodemographics and all-cause mortality in England relative to IMD.
Methods
Experian Mosaic data were merged with IMD quintiles ranked from 1 to 5 (most-to-least deprived) based on full postcodes. Age-standardised mortality rates/100,000

were determined for each of the Mosaic groups and deprivation quintiles, using 2011 census (population) and 2010-2012 mortality data.

Results

Mortality rates in the most-deprived were 1,638 and 1,340 for males and females, compared to 978 and 873 deaths/100,000 in the least-deprived IMD quintile respectively. Mosaic groups showing close correspondence with deprivation were associated with high levels of mortality: Vintage Value [Males=1,434, 95% CI: 1,375-1,493.89; Females=1,098, 95% CI: 1,074-1,122] and Municipal Challenge [Males=914, 95% CI: 864-964; Females=549, 95% CI: 521-578]. However, more affluent Mosaics were associated with lower levels of mortality, even below those reported by IMD: Prestige Positions [Males=231, 95% CI: 218-245); Females=146, 95% CI: 141-151].

Conclusions

Our results showed an overlap between some Mosaic groups and deprivation in patterns of all-cause mortality. These findings provide further evidence on the feasibility of analysing Experian data for health purposes and suggest that further research using these data is justified.

Mitigating the effects of preferentially selected pollution monitoring sites for inference on population health

Joe Watson
University of British Columbia

A general model framework for detecting the preferential sampling of environmental monitors recording an environmental process across space and/or time is presented. The framework models the joint distribution of an environmental process with a site-selection process that considers where and when sites are placed to measure the process. The environmental process may be spatial, temporal or spatio-temporal in nature. By sharing random effects between the two processes, the joint model is able to establish whether site placement was stochastically dependent of the environmental process under study.

Furthermore, if stochastic dependence is identified between the two processes, then inferences about the probability distribution of the spatio-temporal process will change, as will predictions made of the process across space and time. The magnitude of such change can be very large, potentially having big implications for estimating the health effects of many hazardous environmental spatio-temporal processes.

We apply this framework to a case study involving particulate air pollution over the UK where a major reduction in the size of a monitoring network through time occurred. It is demonstrated that a significant response-biased reduction in the air quality monitoring network occurred. We also show that the network was consistently unrepresentative of the levels of particulate matter seen across much of Great Britain throughout the operating life of the network. This may have led to a severe over-reporting of the levels across much of Great Britain. Finally, the potential consequences for health effect estimates are discussed.

What information do we need to model the spatial distribution of zoonotic vectors

William Wint
Environmental Research Group Oxford

What might seem a very simple question actually hides a whole range of complexities. A first step is to establish What type of vector data should be modelled – presence/absence, abundance, trends are just a few of the alternatives. Each has different (and vector dependent) epidemiological relevance, and different requirements for the data needed to train the models. Furthermore, the questions don't concern just the vector data – an extensive suite of covariate datasets are required, and the training data may need enhancing with suitability masks. The modeller also needs to be aware of precisely what mapped outputs are needed. These may well need to be customised to meet the needs of the people who will eventually use the maps to target surveillance, locate risk, or plan control strategies. Modellers may need to derive outputs from the basic models to achieve satisfactory impacts for their outputs. They may also need to translate their outputs into language that are comprehensible to a lay audience. This talk will consider these issues, drawing upon examples of vector modelling at global, continental, regional and local scales to illustrate the challenges modern modellers face to ensure their products are useful in the real world.

3. Abstracts for posters

Simulating spatial distribution of Dengue in Rio de Janeiro using a SIR model coupled to environmental forcing

Leon Diniz Alves
Fundação Getúlio Vargas

Understanding the spatial distribution of dengue is not a trivial task due to the fact that four serotypes of the virus compete between themselves to infect the susceptible population. At the same time, the virus dynamics is linked to that of its vector, Aedes aegypti in Rio de Janeiro. Dengue transmission was simulated from 2010 to 2017 in Rio de Janeiro using a 4-strain SIR model, with cross-immunity, totaling 48 state variables representing all variations of susceptiblerecovered and infectious individuais based on their infection history. The transmissibility and other parameters of the SIR model were made dependent of environment conditions. The environment data used were obtained from public weather satellites data. The simulated dynamics was adjusted to the real dengue incidence data, mathcing seasonal variability that are observed in Dengue incidence including some non-epidemic years Rio de Janeiro.. The next step is to scale up the model to larger regions as a metapopulation model including population mobility. The ultimate goal is to be able to understang the spatiotemporal risk pattern of dengue.

Risk maps, spatial coverage and incentives: An innovative approach for Chagas disease surveillance

Claudia Arévalo-Nieto[1,2], Justin Sheen [3], Luis Vasquez Huerta [2], Ricardo Castillo-Neyra [3], Carlos Condori[1] and Michael Z. Levy[3]
[1]San Agustin University
[2]Peruvian University Cayetano Heredia
[3]University of Pennsylvania

In Arequipa, Peru, entomological surveillance is the final step in a series of efforts aimed to eliminate the Chagas disease vector Triatoma infestans. Here, the complexities of urban infestation and transmission of T. cruzi makes entomological surveillance difficult. In a city of one million inhabitants it is critical to rationally focus surveillance efforts. Toward this end, we have developed VectorPoint, an app that demonstrates risk estimates from an Integrated Nested Laplace Approximations (INLA). The aim of this study was to determine the effect of financial incentives given to vector control personnel, on their use of the risk map, and the spatial coverage of their search. To analyse risk map information use, we used a Cumulative Link Mixed Model (CLMM). We used a metric based on Delaunay triangulation to measure spatial coverage and compared this metric across incentive schemes. In the odds ratio of the incentives fixed effect was 1.05 (95% CI =0.88- 1.25) suggest little effect of the incentives on the use of risk map projections. There was difference between the two arms for spatial coverage (paired t-test, p=0.018) as well as for number of households inspected (paired t-test, p=0.032). These results demonstrate that incentives did not have a differential effect on the risk map information use. However, they had differential effect on spatial coverage achieved by vector control personnel. More complex incentive schemes may be necessary to promote the dual objectives of risk map use and high spatial coverage in the exploration of a city for disease vectors.

Skin cancer geography in Germany

Jobst Augustin
University Medical Center Hamburg-Eppendorf

Skin cancer is the most common type of cancer in Germany. Amon others, the increasing incidence of skin cancer have triggered skin cancer prevention programs. Prevalence and screening utilization shows marked regional differences in Germany. However, reasons and predictors for such differences are still unclear. This contribution is based on two own studies focusing on geographical aspects of skin cancer prevalence and screening frequencies in Germany. Nationwide ambulatory account data (2009-2015) from statutory health insured persons (70.2 million in 2015) in Germany were used. Skin cancer prevalence ((malignant melanoma (MM) and non-melanocytic (NMSC)) and skin cancer screening frequency by dermatologist and primary care physician were determined on a county level. Descriptive- and multivariate analyses were performed to identify spatial patterns in skin cancer screening utilization. Several potential associated factors (sociodemographic and environmental data) were considered. Marked regional variations were observed in prevalences and screening frequency. Multivariate analysis shows e.g. statistically significant positive correlations between higher income and MM/NMSC prevalence. Spatial analysis shows marked clusters with particularly high participation rates in parts of Germany. The multiple regression analyses showed e.g. associations between the skin cancer screening utilization frequency and education. Beside this, significant positive associations between NMSC-prevalence and screening frequency by dermatologists were found. Results show spatio-temporal dynamics in the prevalence of MM and NMSC as well as skin cancer screening frequency in Germany. Sociodemographic indicators have been found as significant influencing factors for both. Smaller spatial levels and a longer period of observation are necessary to make more detailed statements.

Multivariate Space-Time Disease Mapping via Nonparametric Exploration of Disease Risk

Daniel Baer and Andrew Lawson
Medical University of South Carolina

Alzheimer's disease (AD) is a neurological disorder with substantial deleterious effects on cognitive processing, and currently available medications for AD do not clearly alter disease progression. As such, our research focuses on ways in which the progression of AD can be characterized, thereby providing a means for clinicians and public health authorities to hopefully delay or prevent the onset of AD in patients. Specifically, we focus on characterizing how at-risk patients with a precursor to AD, called mild cognitive impairment (MCI), transition from MCI to AD via novel variants of space-time mixture (STM) models we developed in a Bayesian hierarchical model framework. In particular, we make use of a non-parametric quantification of the dependence between estimated MCI and AD risk in spatio-temporal incidence data, called the maximal information coefficient (MIC), which allows us to conditionally share latent temporal disease risk components identified by our model variants based on the estimated dependence between the magnitude of MCI and AD risk over time at the geographic-level. Preliminary results suggest that these novel STM model variants outperform the original STM model in terms of goodness of fit.

Health and mobility: exploring access to primary health care through mobile medical clinics in rural Honduras

Liam Barrett
NHS and Lancaster University

Access to essential health care has been identified as a key priority in achieving the sustainable development goal of good health and wellbeing. This is a challenge in low and medium income countries such as Honduras due to financial constraints and resource limitations creating inequalities in the delivery of health care across the country. In Honduras, whilst 90% of health care is provided by the Ministry of Health, this is predominantly located in urban centres but rural and remote rural areas host to account for 44% of the total population (ref 2016). This situation is recognised by the government's national health model. Nonetheless delivery of care is slow to change and the work of NGOs is crucial in provision of essential health care in such areas. Global Brigades is an international NGO which provides access to health care professionals, medications and education through mobile medical clinics in rural Honduran communities.

In this poster we present a preliminary analysis of the patients' movements between the Global Brigade clinics and how movements are related to the health conditions of the individuals and health accessibility. The poster considers implications of the mobility and immobility of particular populations, asking questions of the ability of those most in need to access care. The national health model, with its focus on social participation is considered in relation to the holistic approach of the Global Brigades aiming to address a communities health and economic needs by empowering community leaders to lead their own way out of poverty through an integrated approach.

Multivariate modelling of geographical differences in three cardiovascular diseases

Kristine Bihrmann
University of Southern Denmark

Acute Myocardial Infarction (AMI), stroke and Atrial Fibrillation (AF) are all cardiovascular diseases and thus may share some common risk factors. Other risk factors may be unique to each disease. The aim of this work was to study geographical differences among the three diseases, since such differences may reflect differences in risk factors. All incident cases (age $\geq$30 years) of AMI, stroke and AF, respectively, in Denmark 2014-2015 were aggregated at municipality level and analyzed using a multivariate Poisson model with a multivariate conditionally autoregressive (MCAR) component to model spatial correlation within disease as well as correlation between diseases. Inference was based on linear models of coregionalization and Markov Chain Monte Carlo methods. Different models were compared using the Deviance Information Criterion. Estimates were adjusted for age and socioeconomic factors. MCAR models provided a better fit to data than separate univariate models, including an IID model assuming no correlation. This indicates correlation both between the three diseases and across municipalities. Preliminary results showed some differences in the geographical distribution of the three diseases. The similarities between the diseases were, however, more obvious, especially for AMI and stroke, which also had the strongest correlation (preliminary results: ρAMI/Stroke=0.8, ρAMI/AF=0.7, ρStroke/AF=0.6). This suggests some common risk factors that were not accounted for, whereas risk factors unique to each disease may be limited. Finding similar patterns in all three diseases strengthens the indication of clustering in the underlying risk of the three considered cardiovascular diseases.

Population mapping and estimation for more targeted vaccination campaigns in the Democratic Republic of Congo

Gianluca Boo
University of Southampton

According to the UN Population Division, the Democratic Republic of Congo (DRC) is the fourth most populous country in Africa. Yet, accurate demographic estimates are currently unavailable because the last nationwide population census was conducted over 30 years ago — in 1984. Uncertainty surrounding these estimates has been shown to be critical for decision making and, among other factors, planning and intervention in vaccination campaigns.

Driven by this public health concern, our work aims to update existing demographic estimates in five provinces of the DRC, including the densely settled Kinshasa province. In doing so, we developed a bottom-up modelling approach to produce gridded population estimates as well as age and gender breakdowns. Our modelling approach incorporates data from over 900 micro-census surveys and various geospatial covariates in a Bayesian hierarchical framework.

The resulting demographic estimates and associated confidence intervals are shared with national and international organizations in the domain of public health through a web-mapping interface. This interface enables to explore and flexibly aggregate the modelling output from the original hectometric resolution to more actionable spatial units, such as customized health zones or administrative units. Our modelling framework is expected to be implemented in other countries with partial or outdated national population censuses.

Methodological considerations in constructing a space-time model of asthma emergency department risk: a case study in South Carolina 1999-2015

Matthew Bozigar
Medical University of South Carolina

Long-term disparities in asthma outcomes are difficult to disentangle using existing data and methods. Limitations in previous research using population-wide hospitalization data aggregated to spatial units for analysis include inadequate or ineffective strategies for handling: spatio-temporal confounding, missing geographic information, variable selection, and expected counts of disease, among other methodological issues. In this research, we address each of these issues by building a space-time model of census tract risk for emergency department (ED) visits due to asthma among South Carolina children from 1999 to 2015 in a Bayesian framework. We control for spatio-temporal confounding by including correlated and uncorrelated spatial, temporal, and spatio-temporal random effects. Employing multiple forms of random effects allowed us to assess risk exceedance probability sensitivity to spatio-temporal extra-variation, and consequently, to stabilize risk exceedance probabilities over time and space. Geographic imputation of missing geographic information was conducted by stochastically imputing census tract identifiers within ZIP code tabulation areas (ZCTA) using areal proportions as probabilities of assignment. We fitted models employing entry parameters with prior distributions on prospective air pollutant, social, environmental confounding, and interaction variables to assist in variable selection. Finally, we assessed sensitivity of variable selection results and coefficient estimates to expected counts calculated using a single average statewide ED visit rate over the study period contrasted with an annually-varying rate. Results showed that controlling for spatio-temporal confounding, imputing missing geographic information, all-at-once variable selection, and using a single statewide rate to calculate expected counts were preferred for accurately detailing asthma risk disparities.

Predicting the TBE and Lyme borreliosis vector Ixodes ricinus in space and time

Katharina Brugger, Melanie Walter, Franz Rubel
University of Veterinary Medicine Vienna

The castor bean tick Ixodes ricinus is the principal vector of viral, bacterial, and protozoan pathogens causing growing public-health issues over the past decades. Most prominent are the tick-borne encephalitis virus and the bacteria complex Borrelia burgdorferi s. l. Proper risk assessments for such tick-borne diseases need quantitative spatial and temporal predictions of host-seeking tick densities. Here, we present a high-resolution density map of host-seeking I. ricinus nymphs compiled for Germany. The input dataset comprises mean annually accumulated nymphal density, as observed by monthly flagging of 100 m2 at 69 sampling sites. With a GLM the observed tick densities were interpolated at unsampled locations by using bioclimatic variables and land cover. While tick densities are very low in urban areas, maximum values up to 1,000 nymphs per 100 m2 are observed in broad-leaved forests. Exemplary one sampling site was selected to answer the question whether it is possible to predict next year's density. With cross-correlation maps and a GLM an eight-year time series of nymphal I. ricinus in Haselmühl (Germany) were analysed. The annual density can be predicted with the mean annual temperature of the previous year, the mean winter temperature (Dec. to Feb.), and the fructification index of the European beech two years prior. Both the year with a low density of 187 nymphs per 100 m2 in 2017 and the extraordinary tick year 2018 with a density of 443 nymphs per 100 m2 were correctly predicted by end of February each year.

MU-MAP (Managing Uncertainties in Modelling Air Pollution)

David Cameron
Centre for Ecology and Hydrology

Here we report on MU-MAP (Managing Uncertainties in Modelling Air Pollution) which was part of the Models to Decisions Network `http://blogs.exeter.ac.uk/models2decisions/`. Our aim was to develop a methodology for assessing uncertainties of small process-based models that are used to estimate depositions of air pollutants to the landscape. Atmospheric deposition affects the health of terrestrial and freshwater ecosystems, and SEPA, a partner in MU-MAP, have the statutory responsibility for air pollution regulation; their decisions depend on the deposition estimated from the CBED (Concentration Based Estimated Deposition) model `http://http://www.pollutantdeposition.ceh.ac.uk/content/cbed-concentration-based-estimated` In MU-MAP we embedded CBED in a Bayesian Hierarchical Model and showed how this facilitated parameter estimation and the estimation of discrepancy due to model structural and process deficiencies. Another goal of MU-MAP was to co-develop with SEPA a web-based tool for mapping and visually communicating the uncertainty that had been quantified.

Spatial inequalities of food retail in the municipality of Rio De Janeiro

Leticia Cardoso, Paulo Cesar Pereira de Castro Junior, Inês Rugani Ribeiro de Castro, Aline Araújo Nobre
Fundação Oswaldo Cruz

The food environment in which people live, work, and study can determine food choices. Inequalities in food availability have been associated with disparity of income in different territories. The study objective was to describe the food environment in the city of Rio de Janeiro, identifying the inequalities in food retail. The addresses of fairs and establishments that market food and meals registered in the municipality up to 2013 were collected and georeferenced. The establishments were classified in three groups: G1 - that predominantly sell foods in natura and / or minimally processed; G2- that sell food from all groups; G3) that predominantly sell ultra-processed foods. Indicators, according to the groups, were constructed to evaluate the food environment: 1) proportion of establishments; 2) density of establishments; 3) ratio of density of establishments. Both kernel maps were created in order to examine the spatial pattern of food retail availability. There was a higher concentration of establishments in higher income areas, independently of the establishment group. A higher proportion of G2 establishments was found, followed by G3 and G1. There was an evident inequality in food availability in the municipality, regions with better income were also those with the highest density of G1, G2 and G3 establishments. The density ratio indicator showed a much greater presence of G3 in relation to G1, especially in the higher income areas. The results point to the need to advance a public supply policy that reorganizes the availability of food in the municipality.

Web data mining: validity of a new tool for food retail data – First results from Rio de Janeiro, Brazil

Leticia Cardoso[1], Mariana Carvalho de Menezes[1], Maria de Fátima de Pina[1], Vanderlei Matos[1], Amélia Augusta de Lima Friche[2], Waleska Teixeira Caiaffa[2], Bruna Costa[2], Milene Pessoa[2], Larissa Mendes[2]
[1]Fundação Oswaldo Cruz
[2]Federal University of Minas Gerais

Considering the challenge of having accurate data on food retail due to common problems related to secondary (difficult to access, compromised quality, data out of date, not georeferenced) and primary data source (costly, time-consuming), we propose an innovative method to get food retail data, through an interface that capture data from earth observation, i.e, Google Earth (GE). The proposed method is relevant to non-commercial use or academic research. The aim of this study is to test the validity of Google Earth data by comparison to direct field observations (gold standard). A secondary aim is to test whether validity differed by type of food outlet and area vulnerability. The scope of the study included the municipality of Rio de Janeiro where 150 census tract area were studied distributed in four strata according to the Health Vulnerability Index. The test method: GE interface is based on the acquisition of health data of interest through services of the Google webservice (in our case, food retail data). We developed a R script to automate digital data acquisition from GE. As Gold Standard we visited all the census tract areas collecting the data from direct field observation. The Results from Rio de Janeiro city will be presented. The sensitivity was 0.87 (CI95% 0.84 – 0.90) and Predictive positive value was 0.76 (CI95% 0.75 – 0.77). The pattern of spatial and social dependence is described, with attention to false positive results. This method can be used to capture the food retail data in Rio de Janeiro and other similar cities in

Latin America.

An analysis of event impact across multiple cancers: Accessing spatio-temporal variation in cancer-specific mortality for Louisiana SEER data

Rachel Carroll
University of North Carolina at Wilmington

Considering the impact of events such as natural disasters or health-related government legislation on disease risk is important. A Bayesian spatio-temporal accelerated failure time model furnishes an ideal situation for modelling events that could impact survival experience via spatial and temporal frailty estimates. Through a hierarchical structure, this model allows the data to detect change-point(s) in addition to generating event-related estimates. A previous study of Louisiana SEER breast cancer data suggested that both Hurricane Katrina and the passing of the Affordable Care Act, colloquially known as Obamacare, had estimable impacts on survival time, 8% shorter and 19% longer survival time in months respectively. For this study, these change-points and event-related estimates were compared across multiple cancers available via Louisiana SEER furnishing an understanding of how event impacts differ in timing and strength by cancer type. The results across multiple cancers showed similarities to trends, both spatial and temporal, found in the original breast cancer dataset. Spatially, differences in survival appeared to be related to socio-economic status, access to health care, access to fresh food, and several environmental pollutants. Temporally, overall trends appeared to differ while similar change points and impacts were estimated for the events of interest. With these results in mind, direct action could be made with the aim of improving survival after detrimental events or in detected spatial areas with worse than average survival as well as maintaining enriched survival after beneficial events or in detected spatial areas with better than average survival.

Detecting spatio-temporal hotspots of scarlet fever in Taiwan with spatio-temporal Gi* statistic

Ta-Chien Chan
Academia Sinica

A resurgence of scarlet fever has caused lots of pediatric infections in East Asia and the United Kingdom. Although scarlet fever in Taiwan has not been a notifiable infectious disease since 2007, the comprehensive national health insurance data can still track its trend. Here, we used data from the open data portal of the Taiwan Centers for Disease Control. The scarlet fever trend was measured by outpatient and hospitalization rates from 2009 to 2017. In order to elucidate the spatio-temporal hotspots, we developed a new method named the spatio-temporal Gi* statistic, and applied joinpoint regression to compute the annual percentage change (APC). The overall APCs in outpatient and hospitalization were 15.1% (95% CI: 10.3%-20.2%) and 7.7% (95%CI: 4.5% -10.9%). The major two infected groups were children aged 5-9 (Outpatient: 0.138 per 1,000 visits; Inpatient: 2.579 per

1,000 visits) and aged 3-4 (Outpatient: 0.084 per 1,000 visits; Inpatient: 1.469 per 1,000 visits). We found the counties in eastern Taiwan and offshore counties had the most hotspots in outpatient setting. In hospitalization, the hotspots mostly occurred in offshore counties close to China. With the help of the spatio-temporal statistic, health workers can set up enhanced laboratory surveillance in those hotspots.

Geostatistical analysis of Malawi's changing malaria transmission from 2010 to 2017

Michael Give Chipeta
University of Oxford

Background
The prevalence of malaria infection in time and space provides important information on the likely sub-national epidemiology of malaria burdens and how this has changed following intervention. Model-Based Geostatitics (MBG) allow national malaria control programmes to leverage multiple data sources to provide predictions of malaria prevalance by district over time. These methods are used to explore Malawi's malaria prevalance changes Malawi from 2010 to 2017.

Methods
Plasmodium falciparum parasite prevalence (PfPR) surveys undertaken in Malawi were assembled. A spatio-temporal geostatistical model was fitted to predict annual malaria risk for children aged 2–10 years at 1x1 km spatial resolutions. Parameter estimation was carried out using Monte-Carlo maximum likelihood methods. Population-adjusted prevalence and populations at risk by district were calculated for 2010 and 2017 to inform malaria control programme priority setting.

Results
2,237 surveys at 1,834 communities were used within MBG framework to predict malaria prevalence. Nationally, there was a 47.2% reduction in mean PfPR2-10 from 29.4% (CI 26.6 to 32.3%) in 2010 to 15.2% (CI 13.3 to 18.0%) in 2017. Declining prevalence was not equal across the country, 25 of 27 districts showed a significant decline ranging from 3.3% to 79% reduction. By 2017, 16% of Malawi's population still lived in areas that support PfPR2-10 ≥ 25%.

Conclusions
Malawi has made substantial progress in reducing the prevalence of malaria over the last seven years. However, Malawi still remains in meso-endemic malaria transmission risk. To sustain gains made and continue reducing transmission further, universal control interventions need to be maintained at a national level.

Bayesian Spatio-temporal SIR model for Epidemic data in Korea

Jungsoon Choi
Hanyang University

Recently, epidemic data in Korea have been collected over space and time and spatio-temporal modeling should be considered to analyze such data. In this work, we analyzed monthly Hepatitis A incidence data in 245 administrative districts of South Korea from 2009-2010 years. Such data also have spatio-temporal dependency structures so we proposed a Bayesian hierarchical spatio-temporal SIR (susceptible-infected-removed) model to estimate the spread distribution of infectious diseases. In order to examine the performance of the proposed model, we considered various fitting and prediction measures. Also, a model without spatio-temporal dependency structure was considered. Finally, we found that considering spatio-temporal dependency structures provides better goodness-of-fit and prediction performance.

Mobile population and risk of malaria in the Amazon Basin

Claudia Torres Codeço, Marcelo Gomes, Leonardo Bastos and Raquel Lana
Fundação Oswaldo Cruz

Mobile populations, including demographic groups that move frequently between different localities due to economic motivations and basic services access, have an important role in the dynamics of malaria in the Amazon region. In order to assess how human exposure at different localities along the year affects malaria probability, we analyse data from a survey conducted in 2015 at Alto Juruá region, in the Brazilian state of Acre. This region is located in one of the largest malaria endemicity pockets in the Americas. A total of 520 households in 40 heterogeneous localities in two municipalities, Rodrigues Alves (RA) and Mâncio Lima (ML), were surveyed. Self-reported malaria case in the last 12 months was used as the response variable in a network-driven hierarchical model. Exposure to each locality is measured by person-time spent in each locality per household, using information of commutation by travel frequency, time per stay, and motivation (study, work, health access, travel). Results indicate that the probability of malaria infection increases with person-hours spent at most rural areas. On the other hand, person-hours spent at Cruzeiro do Sul, which is the closest urban hub, lowers the risk. By taking into account the typical mobility pattern and the estimated effect of each locality, the model indicates that the malaria probability distribution within rural areas has a median of at least 0.26. For rural localities in RA, a quarter of them presented a staggering probability of 0.48 or higher.

Evaluating dengue forecasting model to predict Zika and Chikungunya in Brazil

Flávio Codeço Coelho
Fundação Getulio Vargas

The Mosquito Aedes aegypti is a vector for multiple viruses around the globe. Its distribution is restricted to tropical and subtropical climates, which reflects its sensitivity to temperature, humidity and other weather constraints. The modulation of its life cicle by climate, shapes the seasonality of the diseases it transmits. In Brazil, Aedes aegypti has been mainly associated with the transmission of dengue, making it a marked seasonal disease. In recent years, A. aegypti has also been notably responsible for epidemics of the Zika and Chikungunya virus. In this paper

we explore the performance of dengue forecast models trained on the longer available incidence timeseries to predict the weekly incidence of Zika and Chikungunya as well. We will use a LSTM (long short term memory) recursive neural network model, which we have shown, in a previous work, to yield accurate forecasts for weekly dengue incidence. Climate variables such as temperature, humidity, and atmospheric pressure are also used as predictors. A spatial component built from the incidence at neighboring cities is also included. We present results of the forecast of total incidence of arboviral disease as well as of each disease separately and discuss the relative performances of the model for each of these tasks.

Assessing Spatial Variation in Cognitive Health in Relation to Neighborhood Environment

Anne E Corrigan
Johns Hopkins Bloomberg

It has been shown that modifiable risk factors such as physical activity, sleep quality, and social engagement impact cognitive and functional aging. It is also understood that these factors are impacted by one's physical environment; however, little has been done to characterize the role of the neighborhood as a key influence on long-term trajectories of cognitive health. To protect against the risk of cognitive impairment and Alzheimer's disease, it is essential to improve our understanding of how modifiable risk factors relate to the built environment while simultaneously considering trends in health behaviors and outcomes related to individual characteristics of sex, race, socio-economic status, etc. Using cognitive health data from the National Institute on Aging, which details participants' residential history, socio-demographic status, and mental and physical health since 1980, we examine the relationship between health and place. We specifically assess the association between individual-level cognitive health outcomes and the environment through the use of multivariate regressions. We account for spatial dependence and socioeconomic, demographic characteristics of participants. To characterize neighborhoods we examine variables including walkability, safety, presence of green space, healthy food stores, community centers, etc. Using Moran's I calculations, we assess the spatial dependence of both the health outcomes and the risk factors which are associated with cognitive health. Finally, cluster detection analysis identifies areas of significantly high rates of Alzheimer's Disease and dementia. This identification of areas of high cognitive impairment and characterization of neighborhoods' impact on cognitive aging informs adaptations in health behavior and the built environment.

New guidance for investigating non-infectious disease clusters from potential environmental causes

Helen Crabbe[1], Tony Fletcher[1,4], Rebecca Close[1], Anna Hansell[2], Frederic Piel[3], Rebecca Ghosh[3], Luke Hounsome[3], Mark Reacher[3], Giovanni Leonardi[1,4]
[1] Public Health England
[2] University of Leicester
[3] Imperial College London

[4]London School of Hygiene and Tropical Medicine

Background

Putative clusters of diseases associated with an environmental exposure are often reported to public health agencies. Residents are concerned that local pollution may be linked to apparent similar cancers, chronic diseases, congenital anomalies or other unusual illnesses found in close proximity in either time or space. Public health practitioners need to respond to reports of disease clusters and investigate the cause. Given the complexities of linking environmental exposures to health effects, there is a need for specific guidance on how to best do this.

Aim

National guidelines for addressing non-infectious disease clusters (NIDC) are needed to assist practitioners involved in these investigations. A framework for thorough investigation of any NIDC is proposed.

Methods

The guidance is based on a staged approach with comprehensive steps. At stage 1, a screening process is undertaken to make a decision on whether a cluster should be investigated further. Stage 2 involves the assessment of both the health outcomes and exposure validation, including risk perceptions. Stage 3, if reached, triggers an aetiological investigation with quantitative analysis of the relationship between the health outcomes and the environmental exposure.

Results

Public Health England produced guidelines to assist practitioners in investigating the logical and statistical validity of NIDC and explore linkage to environmental exposures. The framework describes resources that can be utilised to aid cluster investigations such as space-time analysis and GIS.

Conclusions

Investigating potential clusters can be difficult and time-consuming and we propose a systematic, integrated approach for responding to such clusters.

Spatial sampling in the absence of a national census

Édith Darin
University of Southampton

Gridded population data provide a spatially explicit denominator in the study of health-related issues. When the national census data is outdated, geo-statistical models enable the production of gridded population estimates using a bottom-up approach. This approach relies on the collection of survey data across small areas, but the absence of census data challenges the design of representative spatial sampling. To tackle this issue, we propose a spatial sampling design that builds on three elements: 1) contextual strata definition, 2) sample size selection, and 3) spatial sampling. We tested our sampling design in two provinces of the Democratic Republic of Congo, a country that had its last census in 1984. First, we used a principal component analysis combined with a k-means clustering algorithm to create meaningful contextual strata based on the available gridded covariates (e.g., elevation and lights at night). We then estimated the sample size within each contextual stratum by assessing the error in a basic gridded-population model fit using available population estimates. Lastly, we proceeded to stratified spatial random sampling by also enforcing cluster sampling within larger "super" cells. The results of our study show that our proposed spatial sampling design facilitates the selection of a meaningful number of small areas, which reflect the different contextual settings of the study area. In the future studies, we aim at consolidating the three elements of our sampling design by developing a robust comparison of different sampling methods in a simulated environment.

Bayesian spatial modelling and prediction of terrestrial background radiation in Switzerland

Christophe Folly
University of Bern

Introduction
While high and medium doses of ionizing radiation have been shown to increase the risk for cancer, evidence for adverse health implications of low-dose exposures is less clear. Studying the relatively small risks associated with low dose exposures demands large sample sizes, and exposure cannot be assessed with personal dosimetry. Thus, epidemiological studies rely on nationwide exposure models.

Aim
The objective of the study is to develop a Bayesian spatial approach to model and predict the terrestrial background radiation in Switzerland.

Methods
Air born spectrometry measurements of terrestrial background radiation in Switzerland are obtained from the Swiss Nuclear Safety Inspectorate. Data on available covariates include geological maps, land cover data and precipitation. On the first step, we conducted a variable selection using LASSO. We then developed a Bayesian spatial model for the selected variables together with a random effect to account for residual spatial variation. The random effect was assumed to be a realization of a Gaussian field with the Matérn correlation function. Interference was conducted using the Integrated Laplace Approximating (INLA) and the Stochastic Partial Differential Equation (SPDE) approach. To examine the model's accuracy we will use cross validation.

Significance
To our knowledge, this is the first application of the INLA and SPDE framework on terrestrial background radiation. This study will also provide maps of natural background radiation in Switzerland. The predicted maps will be applied to study the effect of low-dose exposure to ionizing radiation on the risk of cancer in children.

Transmission of multidrug resistant tuberculosis in Shanghai: a population-based study

Erjia Ge
University of Toronto

Setting:
Shanghai is a mega city where 39% of the population comprises internal migrants.

Objective:
To examine the different roles played by migrants and permanent residents in the transmission of multidrug-resistant tuberculosis (MDR-TB).
Design:
We conducted a population-based cohort study to assess MDR-TB transmission in Shanghai between 1 January 2009 and 31 December 2012 using genotyping and geospatial analyses.
Results:
A total of 367 MDR-TB cases formed the study cohort. Significant differences between MDR-TB cases who were internal migrants and those who were permanent residents were found with regards to age, sex, region, genetic characteristics and treatment outcomes. Permanent residents had a higher transmission rate than internal migrants (OR 3.36, 95% CI, 1.86-6.09). Permanent residents and genotypic clustering cases had similar clusters in central downtown and some parts of suburban areas. Most of the clusters of internal migrants were found in rural areas bordering suburban areas. Clusters of genotypic non-clustering cases showed patterns that closely matched those of internal migrants suggesting acquired drug resistance in migrants.
Conclusion:
In Shanghai, permanent residents were significantly associated with recent transmission of internal migrants in rural areas were most likely to have contracted MDR-TB through acquired resistance.

Floods and Leptospirosis in Brazilian Municipalities – 2003 to 2013

Renata Gracie, Diego Ricardo Xavier and Roberto Medronho
Oswaldo Cruz Foundation-Fiocruz

In Brazil, there are several cases of leptospirosis related to environmental sanitation problems, especially in flood seasons. Taking into account the issues related to climate change, floods tend to increase. Floods do not homogeneously reach populations; generally, those less fortunate in socioeconomic and environmental sanitation terms are the most affected. In this study, we analyzed leptospirosis in Brazilian municipalities with different population size, floods occurrence, and contextual variables.
Socioeconomic, environmental, and health data at municipal level were located through geoprocessing techniques. We have used Spearman correlation test for contextual variables selection associated to leptospirosis incidence rate and regression tree for exploratory analysis and disease occurrence expected.The greater the number of large flood municipal decrees the higher will be the leptospirosis incidence rates in municipalities. The main predictor variable of leptospirosis incidence rate was the proportion of houses sewage by fosse. The group of 125 municipalities (model node 33) was the analysis units set that presented the highest leptospirosis incidence rates observed – 93.33 cases per 100,000 inhabitants (95% CI 81.08-105.58). The highest leptospirosis incidence rates occurred in municipalities where 31.09% or more of the houses used sewage by fosse and in municipalities with more than 3 flood events. Flood declaration by municipal authority was an important marker of risk for leptospirosis occurrence. The municipalities presenting sanitation problems were in higher risk of leptospirosis occurrence. Regression tree

modeling proved to be very useful to estimate leptospirosis occurrence as well as to mine contextual data related to contextual indicators issues.

Surveillance of the emergence of anthelmintic resistance in human population with soil-transmitted helminths in southern Mozambique

Berta Grau-Pujol
Barcelona Institute for Global Health

Over one billion people worldwide are infected with soil-transmitted helminths (STH), which contribute to malnutrition and delayed cognitive development among others. Currently, the cornerstone for controlling STH is mass drug administration programs with benzimidazole drugs. However, some studies suspect of benzimidazole-resistance, already observed in veterinary medicine in high rates. The study aim was to design and implement a pilot surveillance platform to characterize treatment response and genotypic benzimidazole-resistance in Manhiça district, Mozambique.

We conducted a cross-sectional study with 800 participants older than 5 years old using grid sampling to obtain infection district spatial representativeness. Two stool samples were collected from each participant on two consecutive days. For positive STH infection participants, a third stool sample was collected twenty-one days after treatment. Samples were analysed by Telemann and Kato-Katz techniques. Genotypic benzimidazole-resistance is being evaluated by pyrosequencing to detect single nucleotide polymorphisms related with anthelmintic resistance in the β-tubulin gene. The Kulldorff's scan statistic method was employed for identification of STH infection clustering and benzimidazole-resistance. Bayesian Kriging was used to estimate general spatial trends in infection. 21% of participants were infected with at least one STH. We detected a high rate cluster of infection in Calanga region. The lowest cure rate and egg reduction rate were for T. trichiura (47.3% and 50.2%, respectively). The correlation of β-tubulin SNPs with treatment response will be conferred. Surveillance and better understanding of benzimidazole-resistance mechanisms are needed to monitor its emergence. This could contribute on STH infection control and interruption of its transmission.

Mapping the Effects of Bed Bug Policy in New York City

Kathryn Hacker
University of Pennsylvania

Bed bugs have re-emerged both in the United States and worldwide as a major public health and economic concern in urban areas. While the prevalence of bed bug infestations is high throughout major metropolitan areas there is little consensus regarding policy practices. In this study we use a unique set of Volunteered Geographic Information (VGI) from New York City's 311 database to characterize the bed bug epidemic from 2004 – 2017 and analyse the effects of two major pieces of bed bug legislation. Using this data, we longitudinally analysed the spatial patterns of bed bug complains and confirmed violations across the city and evaluated how the bed bug disclosure laws may have impacted the overall number of complaints and violations. Throughout the study period there was an overall decrease in the number of bed bug complaints ($p < 0.001$) and violation ($p <0.001$) in New York City. Currently we are using the data to construct a spatial-temporal modelling framework to evaluate precisely how the disclosure laws may have influenced this decrease. Using this framework, we will be able to make policy recommendations for other large metropolitan areas.

The Eco-epidemiology of leptospirosis: mapping spill-over infection in complex environments

Kathryn Hacker
University of Pennsylvania

The risk of spill-over infection from animal to human hosts is a complex spatial-temporal process, and is applicable to a variety of infectious disease systems including leptospirosis. Rodent control is largely ineffective at reducing the burden of leptospirosis where Norway rats are the primary reservoir hosts. Delineating the spatial heterogeneity of rat populations in urban slum environments is a key barrier to controlling these pests and leptospirosis. We developed a tracking plate method and geostatistical-modelling framework to quantify the abundance and distribution of rats. We used this method to create high-resolution risk maps for rat abundance and distribution throughout a Brazilian urban slum community where leptospirosis sub-clinical infection is endemic (5.14% CI, 4.10–6.46). We used these predictions as covariates in a spatial-temporal model to predict leptospirosis infection among a cohort of urban slum residents. We captured leptospirosis infections that occurred during the typical dry and rainy seasons from Oct 2014–July 2015. The cumulative incidence of leptospirosis during the two periods was 4.54 (3.40–5.93 CI) and 11.31 (9.52¬–13.31 CI) respectively. The final model mapped leptospirosis risk at fine spatial resolution, and identified clear hotspots of leptospirosis infection. Covariates that were positively associated with leptospirosis infection included rat abundance OR 1.03 (1.01–1.07), cumulative rainfall (m) OR 4.12 (2.45–6.33), and male gender 3.78 (1.96–6.33). These findings emphasize both the importance of spatially explicit design and the incorporation of spatially relevant covariates when working with environmental and biological features. Furthermore, we demonstrate an applicable approach to map spatial heterogeneous processes using the example of urban rats.

Comparing spatio-temporal methods of non-communicable disease surveillance

Mico Hamlyn
Imperial College London

Objective

Non-communicable disease (NCDs) surveillance is less advanced than for infectious diseases. Various methods for spatio-temporal monitoring of NCD are available but limited information is available on their respective overall performance. We set out to compare how various methods perform on simulated small area data.

Methods

Using model simulations in R with grid cells and receiver operating characteristic (ROC) analysis, we compared the sensitivity and specificity of multiple frequentist and Bayesian methods within a multi-dimensional parameter space. We tested for multiple probabilities of a real anomaly being present (base rate); various random effects (RE); and probabilities of an area being randomly given an increased relative risk (RR).

Results

Space-time anomalies were poorly detected by all methods when low RR and low expectations were used. We found that, in most simulations, recent and complex methods (i.e. Bayesian methods) did not outperform simpler frequentist methods (e.g. SaTScan), which were usually easier to set up and quicker to run. Positive predictive values (PPV) tended to be low.

Conclusions

Overall, SaTScan appeared to be the best performing methods in most scenarios. With increasing pressure on public health institutions to tackle the increasing burden of NCDs, careful consideration is required to identify the method best suited to the specificities of the disease being monitored in order to avoid high rates of false positive detection. Despite a low PPV, a routine surveillance system using SaTScan could potentially outperform current surveillance which rely mostly on human reporting.

E-cigarette and Menthol Use Disparities between San Francisco Bay Area Young Adults Living in Cities with and without Flavored Tobacco Sales Restrictions

Louisa M. Holmes[1] and Pamela M. Ling[2]
[1]Binghamton University
[2]University of California San Francisco

Young adults are more likely to smoke menthol cigarettes and e-cigarettes than older adults. Disparities in menthol use are well documented, and e-cigarette use among young people in the United States surged by 78% between 2017 and 2018. Bay Area cities have recently banned flavored tobacco product sales, including menthol in some cases. However, these policies are not uniform across counties. This study uses 2014 SF Bay Area Young Adult Health Survey data, a probabilistic multimode household survey of young adults (18-26) residing in Alameda and San Francisco Counties in California (n=1,363), to first evaluate whether menthol cigarette smoking and e-cigarette use cluster in certain cities or Census block groups, and second determine whether there are differential risk factors for young adults living inside and outside of cities with flavored sales bans. We conducted Getis-Ord Gi* Hotspot Analyses in ArcGIS 10.6.1 and found significant clustering of both menthol and e-cigarette use. Of the block groups with significant clustering, 58% of e-cigarette and 80% of menthol clusters were in areas of the counties that do not have flavor bans. We then performed separate logistic regressions in Stata v15 to com-

pare predictors of use for both products between participants residing in areas with clustering and those outside. Controlling for demographic factors and both self-reported and block-group level aggregate neighborhood characteristics, we found neighborhood disorder to be associated with both e-cigarette (AOR: 1.33, CI: [1.07, 1.65]) and menthol use (AOR: 1.32, CI [1.09, 1.59]) in the hotspot areas.

Random Tree Boosting for Automated Regionalization of Disease Patterns

Mark Janikas[1], Orhun Aydin[1] and Renato Assunção[2]
[1]ESRI
[2]Universidade Federal de Minas Gerais

Regionalization is an essential step in the spatial analysis of public health data. Methodologies that partition areal units are leveraged across a wide spectrum of topics ranging from delineating spatial patterns of disease to planning infrastructure for public health services. Numerically efficient, heuristic regionalization models are known to define sub-optimal regions for small-scale heterogeneity for variables of interest. Exact regionalization methods that search optimal regions are known to be prohibitively slow for a large number of spatial units such as data at the district level. A graph-based heuristic regionalization algorithm, SKATER, is shown to be computationally efficient for big data settings and convergences to optimal results for a large number of regions. We propose a boosting scheme for the SKATER algorithm by searching for consensus among an ensemble of weak regionalizations that are iteratively optimized. The randomized tree search extension to the SKATER algorithm identifies optimal boundaries even when the number of regions desired is small. We compare our approach to frequently used regionalization methods AZP, ARISEL and SKATER in two empirical examples that define planning locations for blood banks and regions of public health indicators.

Malaria resurgence and illegal gold mining activities in the state of Bolivar, Venezuela

Christopher Jarvis
London School of Hygiene & Tropical Medicine

In the preceding decade Venezuela has seen a stark increase in cases of malaria. Although the increase in malaria cases has been considered and analysed at a national level, the change has been mainly driven by increases in the state of Bolivar, where 66% of all cases occurred between 1999 and 2014. One theory for Bolivar being disproportionately affected is due deforestation from illegal gold mining in the area.

Using local point level data, we consider the impact of deforestation on the malaria cases within the Bolivar region over a 10-year period. Counts of malaria were compared with changes in vegetation measured through satellite imagery. The association was assessed using multi-level spatio-temporal models adjusting for climate variables and spatial and temporal dependency. We found strong statistical evidence of a negative association between vegetation cover and malaria cases with areas that had experienced the largest decrease in vegetation reporting the greatest increase in malaria cases over the time period.

Our results are consistent with deforestation in Bolivar being strongly linked with the increase in malaria in the area. As Bolivar is the primary driver behind the increases in malaria in Venezuela stronger focus on regulating mining activities, controlling development and reforestation of areas in Bolivar may help prevent further increases in malaria in the region.

Subnational disparities on immunization coverage and spatial access to immunizing facilities in Kenya

Noel Kanini Joseph[1], Peter Macharia[1], Paul Ouma[1], Victor Alegana[1], Rocco Panciera[2], Maria Muniz[2], Robert Snow[1,3], Emelda Okiro[1], Yaniss Guigoz[4] and Nicolas Ray[4]
[1]KEMRI-Wellcome Trust Research Programme, Nairobi
[2] UNICEF
[3] University of Oxford
[4] University of Geneva

Since 2013, county governments have been responsible for delivery of child immunization services in Kenya and these are routinely delivered at selected and equipped health facilities with few supplemental immunization activities such as campaigns. The contribution of geographic access, defined as distance or travel time to health facilities, to disparities observed in full immunization coverage (FIC) remains poorly defined. The most recent Kenya Demographic Health survey (KDHS-2014) was powered to give estimates of health indicators at county units by survey design and reported coverage of immunization was variable ranging from 32.1%-92.3%. These estimates however, mask heterogeneities within counties hence the need to utilize reliable disaggregated data at finer geographic units. Various factors could explain these inequities and we hypothesize that access to service delivery points plays an important role in utilization of immunization service particularly for regions with inadequate health facilities.

We computed a spatial access metric comprising continuous surfaces (100m by 100m) of travel times to the nearest immunizing health facility and average travel time at each sub-county. Using KDHS-2014 we compute estimates of immunization coverage at sub-county level using small area estimation models and determined the role of spatial access to explain difference in immunization coverage using a multi-level logistic regression model. We adjusted for factors that influenced the pathways of immunization service utilization identified from literature review. From the preliminary analysis, the mean sub-county FIC was 76% with 117 of 295 (40%) sub-counties with FIC <50%. Results from this study will be useful in policy advocacy for spatial access in improving FIC

Bayesian spatio-temporal boundary detection methods for disease mapping

Dayun Kang and Jungsoon Choi
Hanyang University

Most of the studies in disease mapping aim to find risk factors or the spatial and temporal patterns of disease risk. Bayesian hierarchical models with spatio-temporal random effects are most widely used for this purpose and they give smoothing effects over the study areas. However, another important goal of disease mapping studies is to find statistically significant discontinuities in the spatial pattern of disease risk. Since existing Bayesian hierarchical models tend to highly smooth the risk surface, it is hard to detect the boundaries among the spatial units. For this reason, several researches have been conducted regarding boundary detection methods, but they are limited to only spatial analysis. Since spatial boundaries could change over time, spatio-temporal boundary detection methods should be developed. In this study, after a brief review of existing spatial boundary detection methods, we propose spatio-temporal boundary detection models. We illustrate our proposed models using real dataset and simulation.

Bayesian Inference for Crime Incidence in the USA

Nadeen Khaleel
University of Bath

Presence of crime is a constant in cities, however, the relationship between locations of crime incidences and social and economic factors can vary between cities. Modelling crime locations as a point pattern with log-Gaussian Cox process allows us to understand the impact of such factors and the spatial dependency structure within cities but is impeded by sparse and unreliable data. Here we consider algorithms for inference for log-Gaussian Cox processes for more complex models to account for data sparsity and measurement error. We build on an algorithm from Gómez-Rubio & Rue (2018) that combines the flexibility of MCMC and the efficiency of integrated nested Laplace approximations. We can implement this algorithm to perform inference for point patterns across disjoint regions, such as cities that do not share a boundary, and to pool data between these regions to provide more confidence in our inference without a drastic penalty on computation time.

Investigation of the environmental connection between cattle and human EHEC cases

Hyeyoung Kim
National Veterinary Institute

Enterohaemorrhagic Escherichia coli (EHEC) is an important zoonotic infection with clinical symptoms ranging from asymptomatic carriers, mild watery diarrhea to hemorrhagic colitis that sometimes progress with the hemolytic uremic syndrome commonly caused by verocytotoxin producing Escherichia coli O157:H7 (VTEC O157). VTEC is known to be primarily transmitted from an animal reservoir to humans through food, water, the environment. A comprehensive understanding of these transmission routes is essential and significant for improving public health. Researchers have studied the association between human EHEC and cattle VTEC infections revealed that living close to cattle farms can be a significant risk factor for VTEC infection. Most studies, however, have been conducted with data aggregated individual cattle farms or EHEC cases into administrative units. Using aggregated data is a proper way to protect individuals' privacy, but it can cause an issue known as the modifiable areal unit problem and distort the results. In this study, we present work in progress of exploring the relationship between human sporadic EHEC cases and concurrent disease activity in cattle herds in Sweden. We hypothesize that the infectious pressure from the cattle affects the incidence of EHEC human sporadic cases. We use the infectious pressure obtained from the epidemiological cattle herd model in our recent research and the individually geocoded EHEC case data sampled from the residential area that is classified from land use/land cover data. We also investigate the possible risk factors such as population density and weather conditions that can affect the incidence of EHEC cases.

Regional health care and demographic change in Germany

Anne Kis
University Medical Center Hamburg-Eppendorf

Demographic change is one of the main challenges of German and other western societies. The effects are and will be distributed unevenly. In general the share of older people will be growing as the baby boomer generation will reach the age of 65. At the same time the birthrate regresses and young adults move from structural weak regions to urban centers. The therof resultant age-structural shifts lead to regional shifts regarding medical demand on the one hand, but also regarding regional health care on the other hand as demographic change also hits established physicians, who are challenged to find successors. At the moment the so called "adjusted supply rate" is commonly used and the legal basis to reflect the actual health care situation in relation to a given geographical entity. This calculatory variable is meant to ensure nationwide health care in Germany, but studies reveal weaknesses. This study aims to show different ways by means of spatial analyses, in order to measure regional health care situation through the example of dermatological care under special consideration of demographic change. Additionally it provides a method to identify the most vulnerable and therefore considerable regions regarding an assurance of needs-based care. Taking demographic change into account, it is safe to assume that the geographical heterogeneity of dermatological care will increase. This requires greater effort not only in terms of demand planning but also with regard to offering alternativemethods (like telemedicine) of delivering health care and intercommunal cooperation.

Temporal Associations of Determinants on Cardiovascular and Respiratory Emergency Department Visits

Anne Krefis
University Medical Center Hamburg-Eppendorf

Associations between air pollutants, meteorological conditions and respiratory or cardiovascular diseases have been extensively shown. However, few studies have addressed combined effects of air pollutants and meteorological conditions on health. The aim was to investigate associations between meteorological data, data on air pollution and emergency department (ED) visits depending on day of the week, season, and year (2013 - 2015) in Hamburg, Germany. Highly correlated environmental data entered a Principal Components Analysis. By using cross-correlations, time lags with peak correlations of environmental variables on ED visits together with day of the week, year, running day, and season entered a UNIANOVA model.

The final model explained 47% of the variation in respiratory ED visits demonstrating main effects for the day of the week with highest admission rates on Mondays ($B = 10.69$; $p < 0.001$). We observed a time trend showing increasing numbers of respiratory ED visits per each year ($p < 0.001$). The season showed significant effects with lowest ED visits in autumn. Our results did not show direct associations between environmental variables and number of respiratory ED visits. The results for the cardiovascular outcome were less expressive ($R2 = 0.20$). Again, an association between day of the week and number of cardiovascular ED visits was seen ($p < 0.001$). The results suggest that the day of the week had main effects on cardiovascular and respiratory ED visits. In future, we will collect and analyse environmental data on the micro level to achieve higher model quality and better interpretability.

Adjusted, non-Euclidean cluster detection of Vibrio parahaemolyticus in the Chesapeake Bay

Anton Kvit
Johns Hopkins Bloomberg

Vibrio parahaemolyticus (Vp) is a naturally-occurring bacterium found in estuaries, such as the Chesapeake Bay, that can cause vibriosis, a shellfish-borne illness, in humans. Tracking the spatial and temporal distribution of Vp in the Chesapeake Bay, which varies in part due to water temperature, salinity, and other environmental variables, can help identify areas and time periods of high risk. These observations can support interventions used to reduce the burden of vibriosis.

Spatial and spatiotemporal clusters of high Vp abundance were identified among surface water samples in in the Chesapeake Bay between 2007 and 2010. While Euclidean distances between points are often used for cluster detection, non-Euclidean (water) distances were also considered for cluster detection due to the complex nature of the Chesapeake Bay shoreline. Comparison of both methods consistently showed the non-Euclidean cluster detection providing unique and more reasonable clusters than the Euclidean approach. Residuals from univariate and multivariate models were used to identify how clusters changed after controlling for environmental variables. Most clusters tended to decrease in space, time, or significance after adjustment, suggesting these covariates contributed to the original formation of the clusters and as such are useful observation tools for vibriosis risk managers. Clusters that remained after adjustment suggest areas for further study and intervention. These findings reinforce the importance of using non-Euclidean distances when tracking the spatial-temporal variation of Vp as well as the benefits of cluster detection methods for Vp risk management in estuaries.

Onde é o garimpo? A mobile application to map cases of malaria on clandestine gold mining sites in French Guiana

`Yann Lambert`
`Centre d'Investigation Clinique Antilles-Guyane`

The Malakit pilot study experiments the distribution of self-diagnosis and self-treatment kits against Malaria on the borders of French Guiana with Brazil and Suriname to a population of illegal gold miners mostly originating from Brazil. Portuguese-speaking facilitators are trained to interview participants of the study and to collect data using Android tablets with the ODK Collect application after teaching participants how to use a kit. Mapping the mining sites on which participants were likely to be infected with Malaria is essential but challenging since both participants and facilitators are not familiar with a geographical map of French Guiana.

A simple, ad hoc Android application was developed with a database of names and coordinates of known gold mining sites curated by the Parc Amazonien de Guyane and a map of French Guiana displaying the main rivers, villages and cities, and main gold mining areas. The name of a mining site can be entered in a search bar while phonetically matching names are listed dynamically and corresponding mining areas are highlighted in the map. In case no matches are found, several mining areas can be selected manually according to the explanations given by the participant. Once validated by the facilitator, the data is sent back to the pending ODK Collect survey.

At the end of the pilot study, the geographical data collected by Malakit facilitators will be curated and maintained as a reference dataset available for international collaboration in the surveillance and control of Malaria in the region of the Guyana Shield.

Bayesian Spatial Analysis of Drug Mortality Rates in Virginia

`Jong Hyung Lee`
`Virginia Commonwealth University`

Over the last two decades, drug mortality rates increased threefold, from 6.8 to 20.8 deaths per 100,000. Moreover, recent studies highlighted that there is a health disparity in drug mortality rate across the U.S. The main objective of this study was to explore the spatial distribution and pattern of drug mortality rates across Virginia through mortality rate mapping and Bayesian modeling through integrated nested Laplace approximation (INLA). In this study, counties in Virginia (n=133) were the spatial units. We used the observed deaths (based on 2012-2016) for each county to evaluate the mortality rate and its county-specific log relative risks. Among the explanatory variables that indicated statistical significance at the 95% credible interval, an increase of 1 unit in the index of relative rurality variable was associated respectively with an increase of around 140% in the risk of drug mortality rates. For every 1 percent increase in proportion of a locality that had no vehicle, there was 19.0% increase in the risk of drug mortality. For educational attainment variable, an increase of percent in the locality with a Bachelor's degree or higher was associated with a decrease of around 4.7% in the risk of mortality rates. The results of this study provide insights for which specific areas need to be targeted for interventions that will reduce drug mortality rates. The Bayesian models performed in this study can be used to develop new policies in which are more accurate and applicable at the local level.

The role of climate, cities and connectivity in the spread of mosquito-borne diseases in Brazil

Sophie Lee
London School of Hygiene and Tropical Medicine

Mosquito-borne diseases, such as dengue and Zika viruses, pose a high social and economic burden on many tropical and sub-tropical countries. Early vector control aims to reduce the impact of epidemics but relies on accurate modelling tools to identify areas of increased risk and target interventions. Early-warning systems for dengue in Brazil have been developed using Bayesian spatio-temporal models, which produce probabilistic predictions indicating areas at 'high risk' of dengue epidemics. These models account for observed drivers of epidemics, such as climate and socio-economic factors, whilst accounting for unknown latent variables and spatial autocorrelation. Many models use distance as a proxy for connectivity between regions, however the movement of people, goods and services creates links between distant areas. For example, the introduction of dengue into previously naïve areas, such as Acre in the Amazonian region of Brazil, following improvements in infrastructure and accessibility to the state from distant areas with high rates of dengue, suggests that connectivity between regions is not solely defined by their proximity. Therefore, a spatial dependency matrix representative of the flow of humans and vectors around a hierarchical urban network will be necessary to explain this connectivity. In this presentation, I will explore the data sources needed to construct a hierarchical spatial dependency matrix, which will feed into a spatio-temporal Bayesian modelling framework to better understand the role of climate and connectivity in the spread of mosquito-borne diseases in Brazil and to obtain more accurate predictions of future outbreaks.

Detecting Local Clusters of Childhood Malnutrition in the Island Province of Marinduque, Philippines

Novee Lor C. Leyso and Maylin Palatino
University of the Philippines Manila

Background
Under-five malnutrition continues to persist in the island Province of Marinduque, Philippines, with prevalence of some forms of malnutrition even worsening in recent years. Local spatial cluster detection using spatially-referenced data provides a spatial perspective in understanding childhood malnutrition, as key in analyzing patterns of geographic variation, identification of factors associated to place or local characteristics, and for focused targeting of interventions. This study aimed to determine and evaluate spatial clusters of under-five malnutrition across the province and within each municipality at the individual level using household location.
Materials and methods
Data from a province-wide household-based census conducted in 2014–2016 were utilized. Malnutrition was defined as weight-for-age z-score that fall outside the ±2 standard deviations from the median of the WHO reference population. The Kulldorff's elliptical spatial scan statistics using the binomial model was used to locate clusters with high-risk of malnutrition, while controlling for age and socioeconomic status.
Results
Significant cluster of under-five malnutrition was found southwest of the province covering the municipalities of Boac, Buenavista, Gasan, and Torrijos. Except for Municipality of Mogpog, at least one significant cluster of malnutrition was also identified within each municipality—mostly located in coastal areas—suggesting apparent geographical variations across and within municipalities in the province.
Conclusion
Specific programs and interventions should be focused on the identified high-risk clusters to maximize resource. Further studies are also recommended to determine factors affecting variations in childhood malnutrition considering the evidence of spatial clustering found in this study.

Survivable road-trauma resulting in a pre-hospital death: a geospatial examination of access to timely advanced trauma care in New Zealand, 2009-2012

Rebbecca Lilley
University of Otago

Access to trauma care affects severely injured patients' chance of survival. We examined the spatial relationship between geographic locations of road-trauma injuries resulting in a prehospital death and timely access to advanced level trauma care in New Zealand (NZ). Prehospital (on scene/during transfer) road-trauma deaths were identified from NZ's Mortality Collection. Injury survivability was assessed using post-mortem reports. Access to trauma care was estimated using location of road and air ambulance bases, advanced-level hospitals and road network data. Kernel Density Estimation was used to determine geographical clusters of deaths.

Of 294 potentially survivable prehospital road-trauma deaths, 46% did not have theoretical timely access to advanced level hospital care. Continued efforts to improve access to advanced care are required to achieve further reductions in road-trauma mortality.

Tuberculosis and its socioeconomic determinants in Rio de Janeiro: a study applying Geographically Weighted Regression

Mônica de Avelar Figueiredo Mafra Magalhães
Oswaldo Cruz Foundation

The use of GIS and Spatial Statistics take a key role in the analysis of public health data. The incorporation of the spatial dimension in the analysis of tuberculosis can help to understand the dynamics of this disease, since it is a disease with deep social roots, related to poor track conditions. This paper presents a spatial analysis through model Geographically Weighted Regression (GWR) to identify more sites associated with higher incidence of tuberculosis socioeconomic indicators. TB cases were geocoded by address of residence and then quantified by census tract. The Bayesian Local Empirical method was used to eliminate the instability of the incidence rate in the census tracts. A classical multivariate regression model was used and then applied the GWR model with the variables responsible proportion with income between 1 and 2 minimum salary, illiterate proportion, proportion of households with people living alone and average income of the head was used. Compared to the OLS results, the GWR improved understanding of the associations between socioeconomic factors and the incidence of tuberculosis. The coefficient of determination R2 increased from 0.044 to 0.2175 showing greater explanatory power in the spatial model. The maps generated by GWR allow analyzing the distribution of each variable through estimates and their associated t values. Showing that this model is useful in studying the dynamics of tuberculosis and of the local determinants of the incidence rates of the disease.

MUsing nationwide linked microdata to examine the effect of transience on the long-term health conditions and health services utilisation

Lukas Marek, James Greenwell, Matthew Hobbs, Melanie Tomintz, Malcolm Campbell, Simon Kingham and John McCarthy
University of Canterbury

The New Zealand government manages a unique research database of linked microdata about people and households called the Integrated Data Infrastructure (IDI). It consists of data from government agencies, Statistics New Zealand and non-government organisations. The de-identified records in the IDI enable researchers to study human and social interactions on both an aggregated and individual levels. The IDI can also be used determine how home address, and the frequency of address changes (transience) may affect long-term health outcomes and health service utilisation. This study aims to identify the vulnerable transient population

based on geographic transience, social and health service utilisation, and other socioeconomic characteristics of this population. Additionally, this study examines health outcomes for New Zealanders not enrolled in one of regional primary health organisations. Understanding the characteristics of this population together with insights about primary health service utilisation can improve the design of integrated health services and their accessibility for vulnerable populations.

Burn the clipboard: Using a mobile app to bring site evaluation into the digital era

Anna Mölter
University of Manchester

The Urban hybriD models for AiR pollution exposure Assessment (UDARA) study is a collaboration between the University of Manchester and Institut Teknologi Bandung (Indonesia). The UDARA study aims to develop land use regression models to estimate air pollution exposure in Jakarta and a second city in Kalimantan. Exposure estimates will be used in epidemiological analyses utilising data from the Indonesian Family Life Survey.

To develop a land use regression model, data from a number of monitoring sites is required. The monitoring sites should aim to capture the spatial gradients of air pollution within the study area. Since existing spatial gradients tend to be unknown, potential monitoring sites need to be surveyed and surrogate measures of air pollution, such as traffic intensity, need to be evaluated. Traditionally, these surveys have been carried out with paper maps and paper forms. However, advancements in mobile technology now allow for these surveys to be done digitally via mobile apps.

This poster will demonstrate the monitoring site survey app for the UDARA study and present preliminary results for the Jakarta study area. The app was built using Survey123 for ArcGIS, which provides form-centric geo collector apps. It is fully integrated with ArcGIS and apps can be developed and deployed in minutes. The app is easily transferable to other studies. It is particularly useful for the validation of spatial data in geographic regions with a lack of detailed geographic information systems, but which have a critical need for exposure assessment and related epidemiological studies.

Comparison of machine learning and geostatistical models in regions of low malaria risk using a combination of point surveys and aggregated surveillance data

Anita Nandi
University of Oxford

Accurate, high-resolution maps of infection risk are a vital tool for the elimination of malaria in low incidence countries. In these regions, prevalence survey data are sparse and often uninformative due to small sample sizes. Aggregated surveillance data of malaria case counts is becoming more widely available and provides a better measure of low malaria risk than prevalence surveys. However, the surveillance data are often aggregated over large, heterogeneous areas, which limits our ability to learn relationships between the environment and malaria risk. A model that combines point surveys and aggregated surveillance data could leverage the benefits of both. However, such a model must also deal with the fact that these two data types measure different aspects of malaria risk: prevalence and incidence.

We investigate two different methods for combining point surveys and surveillance data. In the first method we train multiple machine learning models on point surveys and then combine the resulting predictions from these with a geostatistical disaggregation model using routine surveillance data. This method is demonstrated for Colombia and Madagascar. The second method is a joint geostatistical disaggregation model that uses separate likelihood functions for the two data types. We illustrate this method for Indonesia, Senegal and Madagascar. In each case we compare the model performance to a baseline model that uses only the aggregated surveillance data. Both methods show an improvement in performance from the baseline model. The model performance using the machine learning predictions outperforms the joint likelihood approach in Madagascar.

Short-term spatial forecasting to aid the elimination of Visceral Leishmaniasis in India

Emily Nightingale
London School of Hygiene & Tropical Medicine

The Indian elimination programme for visceral leishmaniasis (VL) has seen great progress over the past decade. Prompt diagnosis and treatment is critical to continue reducing incidence and avoid epidemics in the increasingly unexposed population. As incidence has fallen, heterogeneity between blocks (subdistricts) has become more apparent, suggesting that targeting resources would improve efficiency. Using national programme data from two endemic states (Bihar and Jharkhand) between 2013-2018, we systematically developed a statistical model of monthly case counts at block level that exploits spatial and temporal correlations in incidence to make short-term forecasts. The model was optimised according to predictive power one month ahead, assessed sequentially through an 18-month test period, and relative trends of several fit and prediction metrics with increasing model complexity were investigated. The final model was dominated by auto-regression (AR) on four distributed lags, followed by contribution of neighbouring blocks (NE) at the same lags. Fitted parameters suggest an overall decreasing trend (AR intercept = 0.80 (SE 0.02), trend = 0.997 (0.0007)) with increasing influence of spatial correlation (NE intercept = 0.027 (0.0018), trend = 1.0065 (0.0017). This produced accurate one-month-ahead predictions, capturing the true count within the middle 50% of the predicted distribution for 86% of test points and within the middle 80% for 95%. Ability to predict varied substantially between blocks; the model was unable to anticipate sudden spikes in incidence. Future work will investigate the accuracy of predictions further ahead in time and develop alternative approaches to capture the epidemic nature of VL incidence.

Spatial models for infants HIV/AIDS incidence using an integrated nested laplace approximation approach

Susan Nzula
Strathmore University

Background
Kenya has made significant progress in the elimination of mother to child transmission of HIV through increasing access to HIV treatment and improving the health and well-being of women and children living with HIV. Despite this progress, broad geographical inequalities in infant HIV outcomes still exist. This study aims to identify areas of inflated HIV risk and associated risk factors for mother to child transmission of HIV.

Methods
Data were obtained from the Early infant diagnosis database that is routinely collected for infants under one year for the year 2017. We fitted hierarchical Poisson models with spatially structured random effects to examine the effects of HAART and breastfeeding on infant HIV risk. Inference was performed using Integrated Nested Laplace Approximation. Spatial random effects were modelled using Conditional autoregressive model. Exceedance probability was used to assess areas where the risk exceeds 1. Model comparison was done using the deviance information criterion.

Results
The National MTCT rate was 3445 per 100,000 live births ranging from 0 in Wajir to 15,385 in Samburu. CAR model performed better in modelling spatial dependency in the data. The results revealed weak evidence that HAART was negatively associated with infant HIV positivity [-0.125, 95% CI = -0.348, -0.102] whereas breastfeeding positively [0.178, 95% CI -0.051, 0.412]. Counties that exhibit risk of HIV above 1 were Turkana, Marsabit, Makueni, Mombasa, Nairobi.

Conclusion
The study provides relevant strategic information required to make investment decisions for targeted high impact interventions to reduce HIV infections among infants in Kenya.

Timely access to emergency services in Brazil

Ricardo Antunes Dantas de Oliveira
Oswaldo Cruz Foundation

Timely use of health services according to need is fundamental to many types of care. Considering emergency services, the timely use means greater possibility to recover from accidents and aggressions. The World Health Organization (WHO) define that all the population should access an Emergency service in no more than two hours. In this study the access of Brazilian population to these services is analysed, using a GIS tool: ACESSMOD, which allows the integration of information about elements that influence the relation between time and distance: population distribution, emergency' services location, roads (including rivers), terrain and land use. 193 emergency hospitals were considered, considering the diversity of services (surgery, rehabilitation), availability of equipment (CT scanners and others) and number of emergency beds. The tool provides an interesting form to evaluate access to care when time is a fundamental aspect. The results show that 74% of Brazilian population has the possibility to reach emergency services in no more than two hours, due to the fact that Brazil's populated is highly concentrated in the Southeast and South regions and around state capitals in the other regions. The problem is that the remaining 26% lives in a vast territory that include the North region, the interior part of the Northeast and Center-West regions and even parts of Southeast. The options to cope with this situation involve providing a better distribution of emergency hospitals, an improvement of services in hospitals not included and even in hospitals considered, where the expansion of emergency beds would be important.

Classifying areas into hot/coldspots and studying the temporal dynamic of these areas within each risk category: A Bayesian two stage method applied to severe malaria in Burkina Faso

Mady Ouédraogo[1,2], Sékou Samadoulougou[2], Alegana VA[3,4], Niko Speybroeck[2] and Fati Kirakoya-Samadoulougou[1]
[1] Université libre de Bruxelles
[2] Université catholique de Louvain
[3] University of Southampton
[4] Flowminder Foundation

Background
Severe malaria is the leading cause of death among children under 5 in Burkina Faso (BF). For policy makers it's important for strategic and preventive purposes to know how the identified malaria hot/coldspots have evolved over time and the identification of areas that, whilst not yet severe malaria hotspots, show a tendency to become hotspots. Here we classified Burkina Faso health districts (operational entity in the country) as severe malaria hot/coldspots, or neither and identified the temporal changes of severe malaria at the district level.
Method
Data on severe malaria cases was extracted from the Burkina Faso District Health Information system covering the period of January 2013 to December 2016. Malaria incidence was estimated by dividing the district aggregated malaria cases by the district age group-specific population. We used a bayesian spatio-temporal negative binomial model to explore the space–time variation in severe malaria incidence. The model was adjusted for climatic and contextual variables. A two-stage classification method has also be performed. Models were implemented in R-INLA and parameters were estimated using MCMC simulation.
Results
The annual incidence of severe malaria increased form 70.0 cases per 1000 (95% CI 69.7–70.3) in 2013 to 74.1 cases per 1000 (95% CI 73.8–74.4) in 2014, followed by a decrease in 2016 of up to 70.1 cases per 1000 (95% CI 69.8–70.4) a decline of over 5%. Amongst the 70 health districts in Burkina Faso, 21(30%) and 33(47%) of them are identified as hotspots and coldspots of severe malaria risk, respectively. Nine of

the hotspots showed a faster increase in severe malaria risk compared to the overall trend while 6 hotspots tended to show low decrease over the period. In addition, 16 health districts not yet identified as hotspots show a much faster increasing trend compared to the overall trend.

Conclusion

Modeling routine data reveals that severe malaria incidence was at critical level in some health districts in Burkina Faso. To prevent health districts from becoming more problematic, understanding the reasons of severe malaria fluctuations by space and time are recommended.

Measuring spatiotemporal dependence of disease using the tau and phi statistics: a literature review, normative commentary and new avenues

Timothy Pollington
University of Warwick

Introduction

The tau and phi statistics typically use geolocation and disease onset times to indicate global clustering and its spatial range. We reviewed works citing the 'seminal' paper by Salje et al (PNAS 2012) or the 'reforms' on the tau by Lessler et al (PLoS ONE May 2016). We highlighted themes and new opportunities for insight into the disease process.

Methods

Using Google Scholar we searched papers or preprints that cited the seminal or reforming papers, and that actively used the statistic rather than mentioning the clustering result of the cited paper. Our extraction protocol produced metadata on the statistic used, study characteristics and results presentation of all the works collected.

Results

Including the seminal and reforming paper, 16 papers actively used the tau, phi or similar statistic. The tau statistic applied to case-only data was the commonest (n=3 papers). Nearly all constructed confidence intervals (CIs) as bootstrapped percentile CIs without consideration as to the shape of the sampling distribution.

Discussion

The range of spatial dependence as indicated by tau is sensitive to the choice of confidence interval due to its shallow gradient at intercept. This is not a trivial inaccuracy considering that these epidemiological works have influence on disease control policy. We present best-practice methods to select and present results from these statistics. Our new tau (rate) statistic better accounts for variable times-at-risk in a study according to individual onset-to-treatment times, and the birth, immunity, migration and death events in the non-cases.

Use of Spatial Analysis to Assess Seasonal Differences of Access to Care in Rural Zambia

Timothy Shields
Johns Hopkins Bloomberg

Many malaria elimination programs in developing countries use reactive case findings to link the diagnosis and treatment of malaria in symptomatic patients presenting at a healthcare facility with testing and treating asymptomatic infected individuals living in the vicinity of an index case household. The premise behind these programs is they are effective in reducing the asymptomatic reservoir and more sustainable than community-wide mass test-and-treat or drug administration campaigns. However, if symptomatic patients are unwilling or unable to travel to a biomedical facility for the initial diagnosis, the program cannot effectively reduce transmission due to missed pockets of asymptomatic carriers. Through interviews conducted at households in rural southern Zambia, we mapped pathways and routes used in the dry and rainy seasons to reach local healthcare providers (biomedical and traditional) in four villages. Additional information on the highest level of use (footpath, bicycle, motorcycle, oxcart or car/truck), and average time needed to reach providers was used to create metrics representing access to biomedical resources providing malaria diagnosis and treatment. Analysis shows variation in the distance to both traditional and biomedical health care resources with longer distances travelled in the rainy season. Tools from the field of spatial statistics were applied, highlighting spatial variation in clustering of these healthcare access metrics within and among the four villages. Results will help identify communities where seasonal and environmental factors serve as barriers to accessing health care facilities in a pre-elimination setting.

District vulnerability to extreme heat temperatures in Mainland Portugal

Susana Pereira da Silva
Nacional de Saúde Doutor Ricardo Jorge

Direct effects of ambient temperature on human health are known and, in the case of high temperatures, can range from moderate skin irritations to the so-called heat stroke. Excess temperature can also have adverse effects on health by aggravating preexisting diseases such as cardiovascular and cerebrovascular diseases, respiratory diseases, diabetes, among others. In Portugal, similarly to other countries, the effects of excess heat on health are known, specifically on mortality, in situations of sudden, intense and prolonged temperature increase, meaning during the so-called heat waves.

The objective of this work was to study the risk of death of mainland Portugal districts to extreme temperatures using summer data (from june 1st to September 30th) between 2008 and 2018. For that a geoadditive Regression model with BayesX was applied to daily mortality rate (/100000 inhabitants) considering as explanatory variables daily maximum temperature, a spatial function, and the weekend day. Almost all districts (15 in 18) had registered temperatures above 39ºC on at least one day, during the period of analysis. The maximum temperature registered over the study period was 47.8ºC on Portalegre (23-07-2015) followed by 46.3º on Santarém (04-08-2018). The extreme heat impacts are not equal all over mainland Portugal as expected since each district has specific temperatures. Considering 20ºC as confort temperature and percentile 99 of district temperatures as adverse temperature the mortality rate ratios varied between 1.14 and 1.28, respectively in Aveiro and Santarém. The higher ratios were found in center and southeast mainland Portugal.

Social and physical geographies of children's neighbourhood mobility: Findings from Neighbourhoods for Active Kids

Melody Smith[1], Victoria Egli[1], Erika Ikeda[2], Jinfeng Zhao[1], Suzanne Mavoa[3], Lisa Mackay[2], Niamh Donnellan[1], Rebecca Amann[1]
[1] University of Auckland
[2] Auckland University of Technology
[3] The University of Melbourne

Children's neighbourhood mobility is an important contributor to health and wellbeing. Neighbourhoods for Active Kids is a cross-sectional study investigating links between children's neighbourhood environments and health (including active travel and independent mobility). Participants were 1102 children aged 8-13 years residing in Auckland, New Zealand, their parents, and school representatives. Children completed an online participatory geographic information systems (GIS) mapping survey. Parent neighbourhood perceptions and reasons for determining their child's mode of travel to school were measured using telephone interviews. School representative perspectives were captured by interviews. Objective measures of the built environment were generated using GIS. Quantitative data were analysed using structural equation and linear mixed modelling, ringmaps, and time-activity diagrams. Qualitative data were analysed using thematic and content coding analysis. Drawing data from multiple sources using mixed methods has provided a robust and comprehensive understanding of factors related to children's neighbourhood mobility.

Colorectal cancer survival in England and Wales by socio-economic status and area deprivation

Charlotte Sturley
University of Leeds

Colorectal cancer (CRC) is the fourth most common cancer and the second most common cause of cancer death in the UK. This study investigates variations in CRC registrations and survival in England and Wales by socio-economic status (SES) and area type. Data were obtained from the Office for National Statistics Longitudinal Study (ONS LS). The ONS LS contains census and life events data, including cancer registrations and deaths, linked between five successive censuses, starting in 1971, for a 1% sample of the population.

14,502 LS members had a diagnosis of CRC. As of 31st December 2016, 20% of these were still alive, 40% had died from CRC and 40% had died from other causes. Exploratory analysis showed CRC registrations and deaths varied by sex, age, SES and cause of death. Registrations were higher among the older age groups, with the greatest number in those aged 70-79 years. The average age at diagnosis and death was higher for females compared to males. Those that died from CRC were younger on average than those that died from other causes.

The study will stratify LS members by different measures of SES (educational attainment, economic activity, housing tenure, car ownership) and by area deprivation (Townsend score). Logistic regression and survival analysis will estimate the probability of death and length of survival (from date of diagnosis to death) by SES and area type. The impact of area trajectories (change in relative deprivation over time) on survival will also be reported.

Association between magnesium in drinking water and cardiovascular mortality

Charlotte Friis Theisen
University of Southern Denmark

The association between magnesium in drinking water and the risk of cardiovascular death has been examined in many studies but never in a Danish context. Some evidence of a protective effect of drinking water rich in magnesium is found in these studies. In the present epidemiological study, register-based data are used along with water samples taken during the past 37 years. The study is designed as a cohort study with a 10-year study period (2005-2014) and includes the entire Danish population aged 30 or more. A Poisson regression model for incidence rates was used to assess the association and included confounders on age, gender, cohabitation and family income as well as adjustment for calendar year. The results showed a significant protective effect of magnesium in drinking water on ischemic heart disease (IHD) and particularly acute myocardial infarct (AMI).The 20% least exposed (≤ 6:65 mg/l), had an increased incidence rate of 24% of dying from AMI compared to the 20% most exposed (> 21:9 mg/l) (IRR=1.24, 95% CI: 1.19-1.29). However, no association was found between the level of magnesium in drinking water and overall cardiovascular death or death from stroke. Further sensitivity analyses have to be carried out to confirm the found association.

Using related parent-child data to examine associations between childhood, parental and environmental characteristics and childhood obesity

Melanie Tomintz [1], Matthew Hobbs, Lukas Marek, John McCarthy, Malcolm Campbell, Simon Kingham
[1]University of Canterbury

Childhood obesity prevalence increased in New Zealand from 8% in 2006/07 to 12% in 2016/17. This is concerning as childhood obesity is associated with adverse physical and mental health outcomes. While some evidence has shown a levelling off in childhood obesity prevalence, this may disguise important ethnic and deprivation-related disparities.

This study is the first simultaneous investigation of child, parental and environmental risk factors and childhood obesity. It uses a unique related dataset with measured height and weight in a nationally representative sample (2013/14 – 2016/17 of nationally representative, pooled, cross-sectional New Zealand Health Survey.

Multilevel logistic first explored risk factors for childhood obesity. A two-step cluster analysis was then used to investigate if child health behaviours clustered and hierarchical cluster analysis accounted for the clustering of environmental features to develop neighbourhood typologies. Finally, an interaction was fitted to test the hypothesis that the environment type would moderate the association between clusters of childhood health behaviours and childhood obesity.

This study provides an important addition to evidence using linked data of 9,026 children and their parents in a unique nationally representative cross-sectional dataset with measured height and weight of both child and parent. This study identified spatial variation and clustering in both childhood behaviours and food and physical activity environment exposures. Both child- and parent-level factors were associated with childhood obesity, however, food and physical activity environment factors were not associated with childhood obesity.

Prevalence and geographical variation of dementia in New Zealand from 2012-2015: utilising big data within the Integrated Data Infrastructure

`Katherine Walesby[1], D Exeter, S Gibb, P Wood, J Starr and T Russ`
`[1]University of Edinburgh`

Introduction
Dementia is a worldwide public health concern. Understanding country-specific burden of disease is required to develop health policies and services. There are no New Zealand (NZ) dementia epidemiological studies. NZ has a big data resource called the Integrated Data Infrastructure system (IDI). Our aims were: (1) Investigate prevalence of late-onset dementia using the IDI, 2012-2015; (2) Compare this with published estimates; and (3) Compare variation between North and South Islands?

Methods
We used a population-based, retrospective cohort design analysing routinely-collected de-identified health/administrative data within the IDI. Our population at risk was all NZ individuals aged ≥60 years. Dementia was defined by ICD-10-AM (Australian Modification) dementia codes or receiving an anti-dementia drug. Dementia sub-types and demographics (age, sex and ethnicity) were also extracted. Counts of dementia rates were calculated for each year, dementia sub-type, and geographical regions. Ethical approval was granted by the University of Auckland.

Results
The number of dementia cases were 11,724 (2012) to 12,528 (2015), 0.27% of NZ's population in 2015 and 1.3% of our NZ cohort ≥60 years. Women and European ethnicity had higher numbers. Around 11% of individuals had more than one dementia diagnosis/sub-type. Dementia counts were higher in North versus South Island.

Conclusions
This is the first study ascertaining dementia rates using NZ data. Our results are lower than published NZ estimates. There are several possible reasons for this. Despite limitations of big data research, our results may more realistically reflect

the differences in NZ.

Spatial clustering of suicide death rate in Kanagawa prefecture, Japan

Kazue Yamaoka[1], Masako Suzuki[1], Keisuke Kuwahara[1], Hirono Ishikawa[1] and Toshiro Tango[12]
[1]Teikyo University
[2]Center for Medical Statistics, Japan

The suicide mortality rate has been high in Japan. Thus, early intervention strategy to solve this problem should be developed in Japan. Spatial scan statistics can detect most likely cluster of suicide deaths and therefore would serve as a useful tool of community evaluation of comprehensive suicide prevention. This study was a spatial epidemiology focusing on the deaths by suicide in Kanagawa Prefecture and examined regional accumulation. We used spatial data of regional statistics about suicide deaths reported by the Cabinet Office from 2011 to 2017.We identified most likely clusters by sex using Flex Scan (Tango and Takahashi, 2005). We compared the results with those from Circular Scan (Kulldorff, 1997). From the results, several clustered regions with the highest occurrence of suicide were detected using Flex Scan, and regional characteristics were clarified. The results using Circular Scan were somewhat different from those of Flex Scan; scanning method using circular spatial scan included lower death rate regions as the most likely cluster. The present results of regional accumulation of suicide in Kanagawa Prefecture would provide important information in policy making for suicide prevention.

Patient's willingness to bypass the nearest physician – Findings from outpatient dermatological care in Germany

Nicole Zander
University Medical Center Hamburg-Eppendorf

Health care planning commonly assumes that patients visit the nearest physician. In reality, however, patients frequently accept greater distances than necessary ("bypassing"). In the area of outpatient care there is only little knowledge about the influence of distance and individual or sociodemographic patient characteristics on the choice of physician. This contribution shows results of two cross-sectional studies and a systematic literature search. The first study was conducted among four dermatologists in Hamburg, Germany. Patients (n=309) with psoriasis and chronic wounds were surveyed about their mobility patterns. Geographical network analyses using exact addresses of both the physicians and the patients were applied to calculate travelled (additional) distances. Predictors for bypassing were education, income, disease severity and mobility impairment. The decision to visit a physician was primarily influenced by the physicians' competence, his range of service and the physician-patient relationship. The study was supplemented by a scoping literature review and a model on determinants for travelled distance in outpatient care was developed. This model is currently validated using a data set of n=497 psoriasis

patients from 29 outpatient care facilities all over Germany. First results indicate that a lower doctor density is one significant predictor for the willingness to accept additional distances and longer waiting times for appointments. The expected results should contribute to the presentation of co-supply effects of metropolitan regions, to determining the influence of individual patient characteristics as well as region-specific infrastructural characteristics on the choice of physician and thus contribute to adequate health care planning.