Strathprints Institutional Repository

Napier, Gary and Lee, Duncan and Robertson, Charles and Lawson, Andrew and Pollock, Kevin (2015) A model to estimate the impact of changes in MMR vaccine uptake on inequalities in measles susceptibility in Scotland. Statistical Methods in Medical Research. ISSN 0962-2802 (In Press) ,

This version is available at http://strathprints.strath.ac.uk/56056/

A model to estimate the impact of changes in MMR vaccine uptake on inequalities in measles susceptibility in Scotland

Gary Napier *   Duncan Lee *   Chris Robertson [†‡§]

Andrew Lawson [¶]   Kevin G Pollock [‡]

November 30, 2015

Abstract

In 1998 an article published by Andrew Wakefield in The Lancet (volume 351, pages 637-641) led to concerns surrounding the safety of the measles, mumps and rubella (MMR) vaccine, by associating it with an increased risk of autism. The paper was later retracted after multiple epidemiological studies failed to find any association, but a substantial decrease in UK vaccination rates was observed in the years following publication. This paper proposes a novel spatio-temporal Bayesian hierarchical model with accompanying software (the `R` package `CARBayesST`) to simultaneously address three key epidemiological

*School of Mathematics and Statistics, University of Glasgow

†Department of Mathematics and Statistics, University of Strathclyde

‡Health Protection Scotland

§International Prevention Research Institute, Lyon, France

¶Department of Public Health Sciences, Medical University of South Carolina

questions about vaccination rates: (i) what impact did the controversy have on the overall temporal trend in vaccination rates in Scotland; (ii) did the magnitude of the spatial inequality in measles susceptibility in Scotland increase due to the MMR vaccination scare; and (iii) are there any covariate effects, such as deprivation, that impacted on measles susceptibility in Scotland. The efficacy of the model is tested by simulation, before being applied to measles susceptibility data in Scotland among a series of cohorts of children who were aged 2-4, in the years 1998 to 2014.

1 Introduction

An article published in 1998 by Andrew Wakefield[1] linked the measles, mumps and rubella (MMR) vaccine with an increased risk of autism. Over the years the article garnered extensive media coverage leading to public concerns surrounding the safety of the vaccination, with MMR vaccination rates decreasing across the UK[2]. By 2003 MMR vaccination rates in Scotland reached a low of 87%, with the corresponding rates in England and Wales dropping below 80%. Repercussions from these decreased vaccination rates were felt in 2013, with large outbreaks of measles occurring in England and Wales[3]. These relatively recent outbreaks are thought to largely be due to unvaccinated young teenagers, i.e. children who were not vaccinated around the time of the MMR vaccination scare[4]. Other studies were conducted to investigate the potential link between MMR vaccination and autism around the time of the Wakefield[1] study, and Gilbert and Hiejbel[5] and Taylor et al[6] found no such associations. For more information on a summary of studies failing to find any association linking the MMR vaccine with autism see El-

liman and Bedford[7]. As a result of this evidence the Wakefield article[1] was partially retracted in 2004, before being fully discredited in 2010.

In this paper, we investigate spatial and temporal trends in estimated MMR susceptibility in Scotland for pre-school children (aged 2-4) between 1998 and 2014. Our data consist of the estimated numbers of children susceptible to measles in each intermediate geography in Scotland, which are small geographical units containing 4000 residents on average. Susceptibility to measles is based upon the receipt of one or two vaccinations, with each having a 10% failure rate[8]. The aims of this paper are three-fold. Firstly, we aim to estimate the magnitude of the increased measles susceptibility associated with the MMR vaccination scare linked to the Wakefield article[1], and assess whether susceptibility has decreased in recent years. Secondly, we assess whether the magnitude of the inequality, as measured by the spatial variability, in measles susceptibility in Scotland increased at the same time and whether spatial variation has subsequently decreased. Thirdly, we determine whether any area based covariates, such as deprivation, have any impact on measles susceptibility in Scotland. Addressing these questions will allow us to identify areas which are at the greatest risk of measles outbreaks, and to quantify the changing nature of measles health inequalities between affluent and deprived communities in Scotland.

We answer the questions posed by developing a novel spatio-temporal Bayesian hierarchical binomial logistic regression model, with inference based on Markov chain Monte Carlo (MCMC) simulation. The spatial patterns in measles susceptibility are captured by random effects represented by the general class of Gaussian Markov Random Field (GMRF[9]) prior distributions. A special case of GMRF priors called conditional autoregressive

(CAR[10]) priors are placed on a set of random effects used to capture the spatial pattern in measles susceptibility between the intermediate geographies in Scotland, while a first-order random walk process is placed upon a second set of random effects to capture the overall temporal trend. The model used in our analysis can be implemented using the `CARBayesST` package, which is part of the statistical programming language `R`[11].

The paper is organised as follows: The measles susceptibility data are presented in Section 2, together with exploratory analysis. The Bayesian hierarchical model is described in Section 3, with a simulation study to illustrate model performance given in Section 4. Results obtained from the measles susceptibility case study are given in Section 5, with concluding remarks provided in Section 6.

2 Case study - Measles susceptibility in Scotland

2.1 Data description

Since 1988 in Scotland, when the MMR vaccination was introduced, individual records of all children in Scotland are collected by the National Health Service (NHS) boards and kept in the Scottish Immunisation & Recall System (SIRS) by the Information and Statistics Division (ISD). This holds, dates of birth, dates of vaccination and postcode of residence (http://www.isdscotland.org/Health-Topics/Child-Health/Child-Health-Programme/Scottish-Immunisation-Recall-System.asp). The data analysed here consist of the estimated numbers of children elegible to attend pre-school (aged 2-4) from non-overlapping two year birth cohorts who

are susceptible to measles (Y_{kt}), and the total number of pre-school children (N_{kt}) in the 1235 (indexed by k) intermediate geographies in Scotland between 1998 and 2014 (indexed by t for every 2 years). For example, the children eligible to attend pre-school in September 1998 were born between March 1994 and February 1996, while children eligible to attend pre-school in September 2000 were born between March 1996 and February 1998, and so on for consecutive non-overlapping two year birth cohorts. From (Y_{kt}, N_{kt}) emerges $\hat{\theta}_{kt} = Y_{kt}/N_{kt}$, the estimated proportion of children who are susceptible to measles, where measles susceptibility is based upon the receipt of one or two vaccinations that each have a failure rate of 10%. For this study a child is deemed completely susceptible if they have not had any vaccinations.

Data on potential explanatory variables are obtained from the Scottish Neighbourhood Statistics website (http://www.sns.gov.uk), to estimate any covariate effects that may impact on the spatial variation in measles susceptibility in Scotland. The median house price (MHP), and the percentage of the working age population claiming jobseeker's allowance (JSA) in each intermediate geography, are used as measures of socio-economic deprivation, where the latter is a benefit paid to working age people who are unemployed. The MHP spans the period from 1998 to 2012, so the MHP in 2012 is also used here for 2014. For the JSA, the median over four quarterly data points is computed for each year, where due to issues of data availability the JSA for 2001 is used for the years 1998 and 2000, while the JSA in 2009 is used for the years 2010, 2012 and 2014. A total of 33 intermediate geographies contained missing JSA percentages for certain years, and were replaced by the percentages, in the same area, from the closest year containing data. The percentage of pupils in primary and secondary schools from ethnic mi-

nority groups is used as a measure of ethnicity (EM), and is available from 2004 to 2012. The percentages of pupils in ethnic minority groups in 2004 is used for all preceeding years, while the percentages in 2012 are used in 2014. A total of 10 intermediate geographies in 2006 contained missing data, with those values replaced with the percentages, in the same area, from the previous year. To remove temporal trends in these variables and to assess the spatial association between each covariate and the response, each of these covariates have been mean centred for each year.

Figure 1 displays boxplots of the raw proportions of pre-school children susceptible to measles in Scotland between 1998 and 2014. A gradual steady increase in the proportion is observed between 1998 and 2004 where susceptibility peaks, before decreasing in successive years. Figure 1 also displays the relationship between the proportion susceptible and each of the covariates, where the solid lines are obtained using locally weighted scatterplot smoothing (LOWESS[12]). Each of the covariates appear to have non-linear relationships with the proportion of children susceptible to measles, exhibiting generally increasing proportions for increasing median house price, percentage of working age population claiming JSA, and the percentage of pupils in ethnic minority groups.
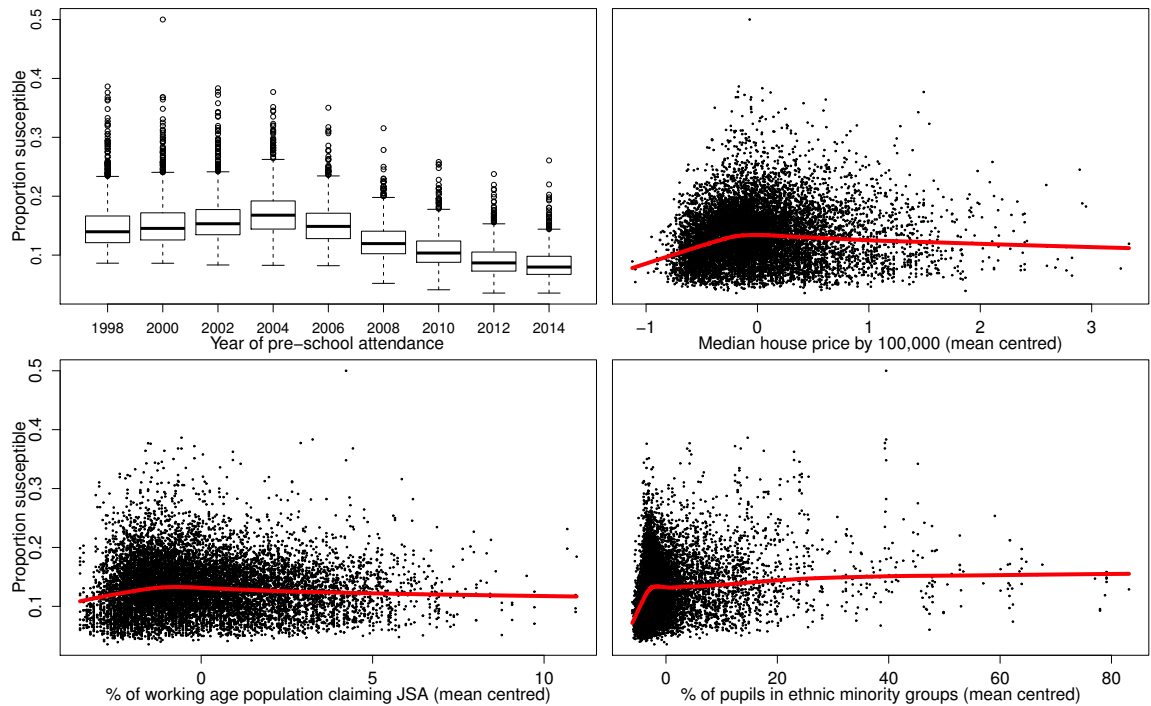
Figure 1: Boxplots of the raw proportions of pre-school children susceptible to measles by year, and scatterplots of their relationships with each of the mean centred covariates. The solid lines are obtained from locally weighted scatterplot smoothing (LOWESS).

Figure 2 displays maps of the raw proportions of children susceptible to measles in Scotland attending pre-school in 1998, 2004, 2008 and 2014. These four years are selected as measles susceptibility rises from 1998 and peaks in 2004, before decreasing again. The most noticeable increases in measles susceptibility are in the less densely populated western and northern regions of Scotland. From 2004 onwards measles susceptibility then decreases, with susceptibility rates at their lowest in 2014.
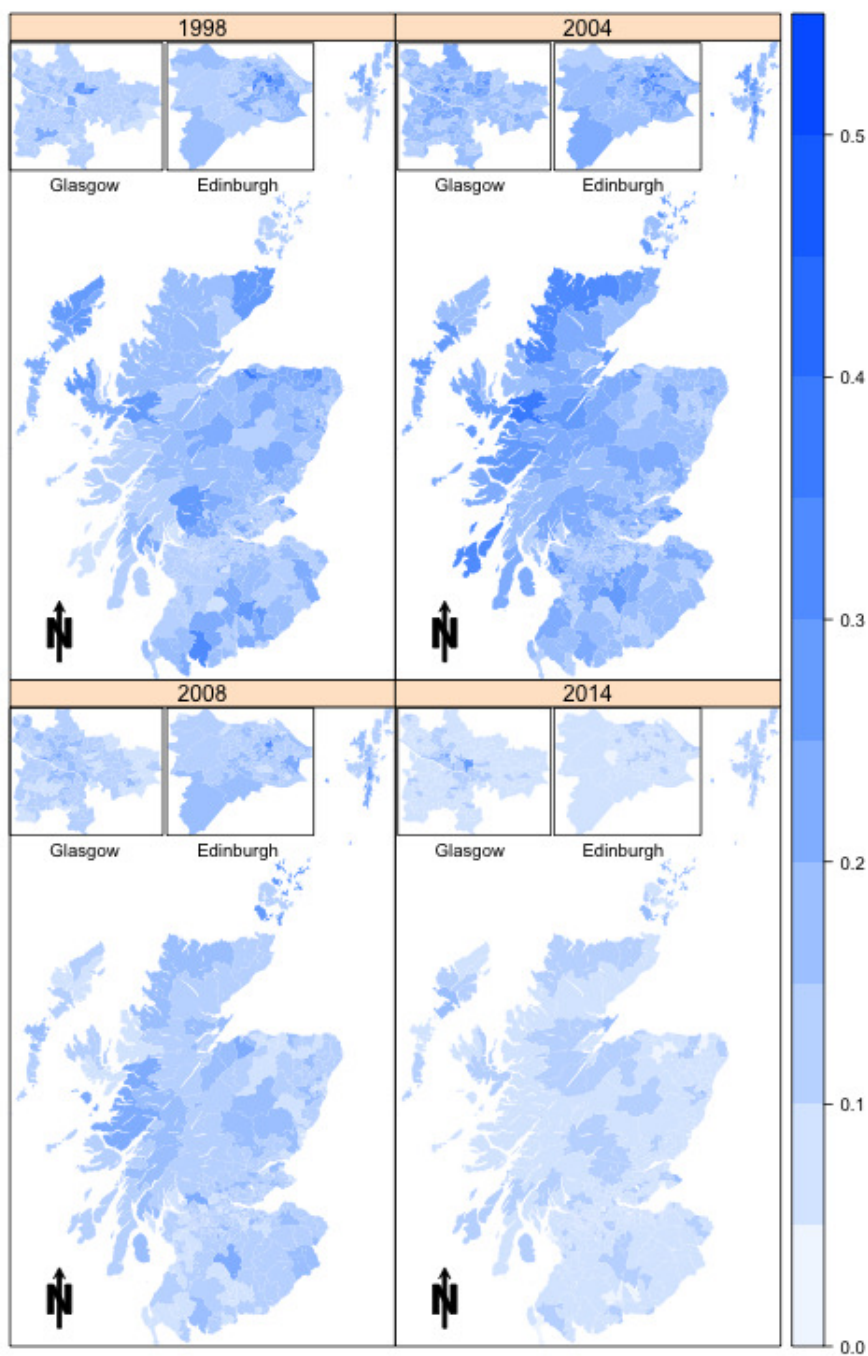
Figure 2: Raw proportions of children susceptible to measles in Scotland attending pre-school in 1998, 2004, 2008 and 2014. The panels in the top-left corner highlight the cities of Glasgow and Edinburgh in the centre of Scotland.

2.2 Exploratory analysis

To assess the appropriateness of using spatio-temporal autocorrelation models, we assess the presence of spatio-temporal autocorrelation remaining in the data after accounting for the known covariates. We fit the following binomial generalised linear model to the data:

$$Y_{kt} \sim \text{Binomial}(N_{kt}, \theta_{kt}),$$
$$\ln\left(\frac{\theta_{kt}}{1-\theta_{kt}}\right) = \beta_0 + S_1(\text{MHP}_{kt}) + S_2(\text{JSA}_{kt}) + S_3(\text{EM}_{kt}), \quad (1)$$

where $S_i(\cdot) = \sum_{j=1}^{3} B_i(\cdot)\beta_{ij}$ for $i = 1, 2, 3$, is a natural cubic spline with basis functions $B_i(\cdot)$ of each covariate with 3 degrees of freedom to allow for the non-linear relationships observed in Figure 1. The degrees of freedom was chosen to be 3 as it provides a low order smooth shape. Here recall that Y_{kt} and N_{kt} denote the number of children susceptible to measles, and the total number of children in intermediate geography k in time period t, respectively. Spatial autocorrelation in the residuals was assessed by computing Moran's I[13] statistics and performing permutation tests on the residuals from Model (1) separately for each year. The values of the Moran's I statistics obtained ranged between 0.12 and 0.38, with the statistics generated from 10 000 random permutations of the data yielding p-values of less than 0.0001 in all cases, thus indicating significant spatial autocorrelation. The temporal autocorrelation at lag 1 for each intermediate geography was then computed from the residuals of Model (1), and the mean value across all intermediate geographies was 0.04, which would indicate a weak but positive temporal autocorrelation structure in the data. The next section details the Bayesian hierarchical model proposed to model this spatio-temporal structure observed in the data to address the questions outlined in Section 1.

3 Methods

Our study region of Scotland is partitioned into $k = 1, \ldots, K = 1\,235$ areal units defined by intermediate geographies, with data on measles susceptibility across $t = 1, \ldots, N = 9$ time periods. A Bayesian hierarchical binomial logistic regression model is proposed for modelling the proportion of children susceptible to measles. Denote by θ_{kt} the probability of a child being susceptible to measles in intermediate geography k in time period t, then the first level of our Bayesian hierarchical model is as follows:

$$Y_{kt} | N_{kt}, \theta_{kt} \sim \text{Binomial}\left(N_{kt}, \theta_{kt}\right),$$
$$\ln\left(\frac{\theta_{kt}}{1 - \theta_{kt}}\right) = \mathbf{x}_{kt}^{\top}\boldsymbol{\beta} + \phi_{kt} + \delta_t. \tag{2}$$

The logit probabilities of susceptibility to measles are modelled as a linear combination of a $p \times 1$ vector $\mathbf{x}_{kt}$ of covariates (including natural cubic spline basis functions) and their respective regression parameters $\boldsymbol{\beta}$, and two random effects ϕ_{kt} and δ_t. A diffuse multivariate normal prior distribution is placed upon the regression parameters $\boldsymbol{\beta}$, that is $\boldsymbol{\beta} \sim \text{N}(\mathbf{0}, 1000\mathbb{I})$, where $\mathbb{I}_{p \times p}$ is the identity matrix. The random effects ϕ_{kt} and δ_t are included to model the spatial and temporal trends and autocorrelation in the data once the covariate effects have been accounted for. Firstly, the temporal random effects $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_N)$, have a random walk of order 1 prior distribution given as follows:

$$\delta_t \sim \text{N}\left(\delta_{t-1}, \sigma^2\right), \quad \text{for } t = 2, \ldots, N,$$
$$\sigma^2 \sim \text{Inverse-Gamma}\left(0.001, 0.001\right). \tag{3}$$

The mean of the temporal random effect δ_t at time period t is the effect at the previous time period δ_{t-1}, where for the first time period $f(\delta_1) \propto 1$. The temporal random effects thus capture the overall temporal trend in measles susceptibility across Scotland. Alternatives to a first-order random walk include a changepoint model, a second-order random walk and splines, and further details of trend models are given by Ruppert et al[14] and Woodward et al[15].

Let $\boldsymbol{\phi}_t = (\phi_{1t}, \ldots, \phi_{Kt})$ denote the vector of all of the spatial random effects at time period t, then this represents the residual spatial structure in susceptibility in year t after adjusting for the covariate effects. A multivariate normal distribution is placed on $\boldsymbol{\phi}_t$:

$$\boldsymbol{\phi}_t \sim \mathrm{N}(\mathbf{0}, \tau_t^2 \mathbf{Q}(\mathbf{W}, \rho)^{-1}) \ \ \text{for } t = 1, \ldots, N. \tag{4}$$

This spatial random effects model was proposed by Leroux et al[16], and are zero mean centred, with the spatial autocorrelation in the data being modelled by the precision matrix $\mathbf{Q}(\mathbf{W}, \rho) = [\rho\,(\mathrm{diag}(\mathbf{W1}) - \mathbf{W}) + (1 - \rho)\mathbb{I}]$, where $\mathbf{1}$ is a $K \times 1$ vector of 1's so that $\mathrm{diag}(\mathbf{W1})$ is a diagonal matrix with diagonal elements equal to the row sums of $\mathbf{W}$. The matrices $\mathbf{W}$ and $\mathbb{I}_{K \times K}$ are the $K \times K$ neighbourhood matrix and identity matrix, respectively. For the neighbourhood matrix $\mathbf{W}$ the jk-th element $w_{jk} = 1$ if areas j and k are defined to be neighbours. Here areas are defined to be neighbours if they share a common border, otherwise $w_{jk} = 0$, and $w_{kk} = 0$ for all k. The precision matrix $\mathbf{Q}(\mathbf{W}, \rho)$ is a weighted average of spatially dependent and independent correlation structures, and the spatial dependence is clear to

see from the full conditional specification:

$$\phi_{kt}|\boldsymbol{\phi}_{-kt}, \mathbf{W}, \rho, \tau_t^2 \sim \mathrm{N}\left(\frac{\rho\sum_{j=1}^{K}w_{kj}\phi_{jt}}{\rho\sum_{j=1}^{K}w_{kj}+1-\rho}, \frac{\tau_t^2}{\rho\sum_{j=1}^{K}w_{kj}+1-\rho}\right), \quad (5)$$

where $\boldsymbol{\phi}_{-kt} = (\phi_{1t}, \ldots, \phi_{k-1t}, \phi_{k+1t}, \ldots, \phi_{Kt})$. A Uniform$(0,1)$ prior distribution is placed upon the dependence parameter ρ, which governs the strength of the spatial autocorrelation. When the spatial autocorrelation is strong such that $\rho = 1$ the full conditionals in (5) reduce to the Intrinsic CAR model proposed by Besag et al[10], while if $\rho = 0$ the random effects become spatially independent with zero mean and a common variance. To ensure identifiability in the model each set of $\boldsymbol{\phi}_t$ are zero-mean centred at each iteration of the MCMC algorithm, and thus τ_t^2 captures the spatial variation at time t around the overall trend δ_t.

The amount of spatial variation between the random effects at time t is controlled by the parameter τ_t^2, which we allow to differ across time periods, and place independent non-informative inverse-gamma priors on each τ_t^2:

$$\tau_t^2 \sim \text{Inverse-Gamma}(0.001, 0.001) \ \ \text{for } t = 1, \ldots, N. \quad (6)$$

By allowing the parameter τ_t^2 to change temporally we are able to estimate the changing nature of spatial variability and hence the magnitude of inequalities in measles susceptibility across Scotland at different time periods. Note, we keep ρ constant across time as this ensures that the $(\tau_1^2, \ldots, \tau_N^2)$ are comparable as ρ and τ_t^2 are not independent. To sample from the posterior distributions of the parameters in the model MCMC methods involving Gibbs sampling and Metropolis-Hastings moves are used. Software to implement the model is provided in the R package `CARBayesST`.

4 Simulation study

A simulation study is conducted to assess the performance of the model outlined in Section 3, particularly its ability to correctly estimate the changing levels of spatial variability and temporal trend.

4.1 Data generation

Similar data to the measles susceptibility data are generated from Model (2) for $N = 9$ time periods. One hundred datasets are simulated under eight different scenarios, which differ in the temporal trend in a) the level of spatial variability and b) the overall average susceptibility. In both cases four different temporal trend scenarios are considered: constant, increasing, decreasing, and fluctuating, where the spatial variability at time t is denoted by τ_t^2 while the average (over space) susceptibility is $\bar{\theta}_t = \frac{\exp(\beta_0 + \delta_t)}{1 + \exp(\beta_0 + \delta_t)}$. Data were generated under these eight scenarios with and without covariates, and also under the presence of moderate spatial autocorrelation ($\rho = 0.7$), to assess their impact on model performance. The results from our simulations observing changing spatial variability and temporal trend are provided below. In all cases the model described in Section 3 was applied to each simulated data set, where inference was based on 1 000 independent samples obtained from generating 100 000 samples, where the first 50 000 were treated as the burn-in period and the remaining 50 000 were thinned, where every 50th draw was stored and the remainder discarded.

4.2 Results - changing spatial variability τ_t^2

Figure 3 displays boxplots of the posterior medians of the spatial variance parameters τ_t^2 from the 100 simulated datasets, and these results relate to simulated data with moderate spatial autocorrelation ($\rho = 0.7$) and no covariates. From Figure 3 it is clear that the model is able to accurately estimate the true temporally varying spatial variation τ_t^2 used to generate the data, as the true values indicated by the points, are in the centre of each of the boxplots. In terms of uncertainty quantification, Table 1 displays the coverage rates by year obtained from the simulation study with spatial autocorrelation $\rho = 0.7$ and no covariates, and indicates correct uncertainty quantification under each scenario. The table also shows that the estimated variances τ_t^2 are largely unbiased. Similar results were also obtained from the simulated data sets generated under the presence of no spatial correlation ($\rho = 0$) and strong spatial autocorrelation ($\rho = 0.99$) and with covariates, and these results are not shown for brevity.

Table 1: Bias (to 4 d.p.) and coverage (%) from the posterior distributions of τ_t^2 obtained from the 100 data sets simulated under moderate spatial autocorrelation ($\rho = 0.7$) and no covariates to assess changing spatial variability under four different scenarios: (i) constant spatial variability, (ii) increasing spatial variability, (iii) decreasing spatial variability, and (iv) fluctuating spatial variability.

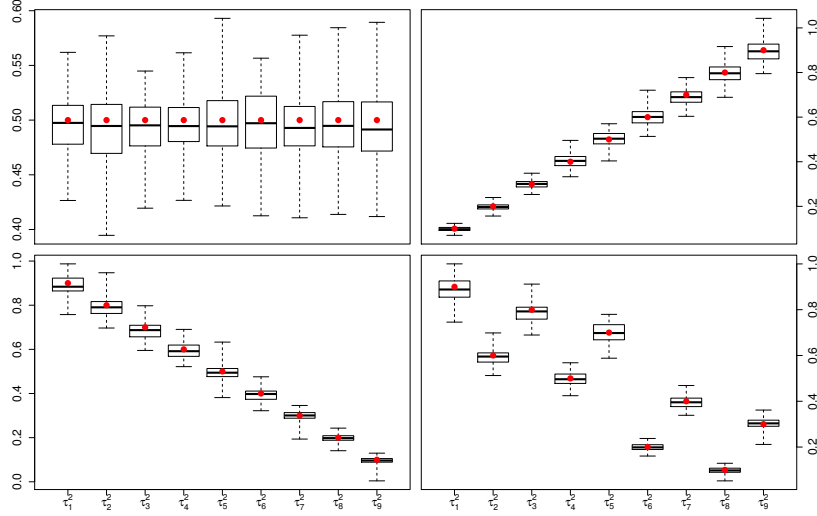| Scenario	Spatial variability								
	τ_1^2	τ_2^2	τ_3^2	τ_4^2	τ_5^2	τ_6^2	τ_7^2	τ_8^2	τ_9^2
(i)	-0.0019	-0.0066	-0.0070	-0.0021	-0.0013	-0.0022	-0.0053	-0.0033	-0.0045
	95%	91%	95%	97%	95%	97%	93%	94%	89%
(ii)	-0.0016	-0.0027	-0.0005	0.0027	0.0033	0.0012	-0.0079	-0.0051	-0.0014
	94%	93%	97%	94%	96%	97%	97%	96%	92%
(iii)	-0.0095	-0.0058	-0.0138	-0.0031	-0.0053	-0.0036	-0.0009	-0.0013	-0.0045
	93%	97%	96%	92%	93%	93%	96%	91%	90%
(iv)	-0.0090	-0.0047	-0.0078	-0.0013	-0.0009	0.0007	-0.0036	-0.0004	0.0040
	93%	95%	94%	94%	94%	98%	97%	91%	96%

14

Figure 3: Boxplots of the 100 posterior medians of τ_t^2 obtained from the 100 data sets simulated under spatial autocorrelation $\rho = 0.7$ and no covariates to assess changing spatial variability under four different scenarios: (i) constant spatial variability (top left panel), (ii) increasing spatial variability (top right panel), (iii) decreasing spatial variability (bottom left panel), and (iv) fluctuating spatial variability (bottom right panel). The points correspond to the true values used in the simulation study.

4.2.1 Changing temporal trend δ_t

Figure 4 displays boxplots of the posterior medians of the average (over space) temporal trend $\bar{\theta}_t = \exp(\beta_0 + \delta_t)/1 + \exp(\beta_0 + \delta_t)$ over the 100 simulated datasets, and these results again relate to data generated with moderate spatial autocorrelation ($\rho = 0.7$) and no covariates. From Figure 4 it is clear that the model is able to capture the true temporal trend in the data for each of the four different scenarios, as the true values (the dots) are close to the centre of each of the boxplots of simulations. Coverage rates from the study are displayed in Table 2 and again suggest appropriate uncertainty quantification from our model under each scenario. As before the table shows the estimates are largely unbiased. Similar temporal trend

results were obtained under the presence of no spatial correlation ($\rho = 0$) and strong spatial autocorrelation ($\rho = 0.99$) and with covariates, but are again not shown for brevity.

Table 2: Bias (to 4 d.p.) and coverage (%) from the posterior distributions of $\bar{\theta}_t$ obtained from the 100 data sets simulated under moderate spatial autocorrelation ($\rho = 0.7$) and no covariates to assess changing temporal trend under four different scenarios: (i) constant temporal trend, (ii) increasing temporal trend, (iii) decreasing temporal trend, and (iv) fluctuating temporal trend.

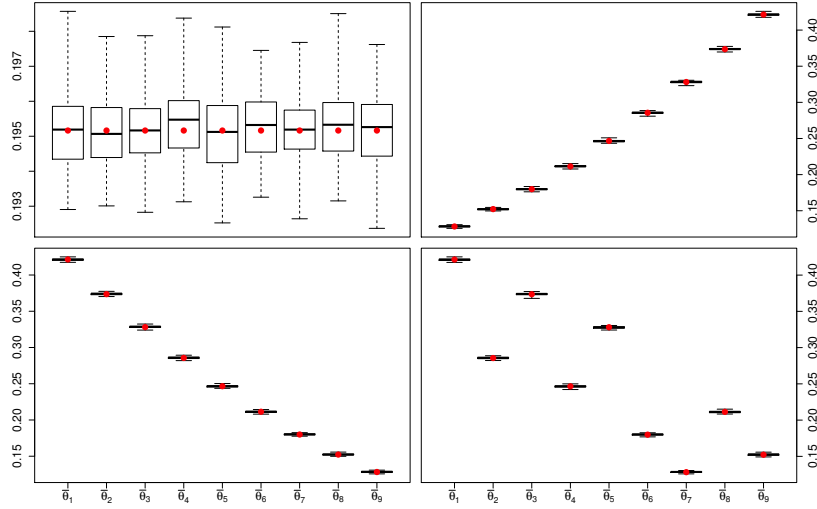| Scenario	Temporal trend								
	$\bar{\theta}_1$	$\bar{\theta}_2$	$\bar{\theta}_3$	$\bar{\theta}_4$	$\bar{\theta}_5$	$\bar{\theta}_6$	$\bar{\theta}_7$	$\bar{\theta}_8$	$\bar{\theta}_9$
(i)	0.0001	0.0001	0.0000	0.0003	0.0000	0.0001	0.0002	0.0001	-0.0001
	94%	95%	98%	95%	96%	98%	97%	95%	98%
(ii)	0.0001	-0.0001	-0.0001	0.0003	-0.0001	-0.0001	-0.0001	0.0001	-0.0001
	96%	92%	95%	92%	93%	95%	95%	97%	96%
(iii)	-0.0002	0.0000	0.0002	0.0000	-0.0002	0.0000	0.0002	0.0000	0.0001
	99%	94%	95%	95%	95%	96%	96%	89%	91%
(iv)	-0.0001	0.0002	-0.0001	0.0001	-0.0003	0.0002	0.0002	0.0002	0.0000
	99%	98%	95%	96%	95%	96%	94%	92%	92%

Figure 4: Boxplots of the 100 posterior medians of $\bar{\theta}_t$ obtained from the 100 data sets simulated under spatial autocorrelation $\rho = 0.7$ and no covariates to assess changing temporal trend under four different scenarios: (i) constant temporal trend (top left panel), (ii) increasing temporal trend (top right panel), (iii) decreasing temporal trend (bottom left panel), and (iv) fluctuating temporal trend (bottom right). The points correspond to the true values used in the simulation study.

5 Measles susceptibility results

The results in this section are based on the model outlined in Section 3, where 150 000 samples were generated, and following a burn-in period of 50 000 samples convergence was assessed to have been reached using Geweke[17] statistics. Subsequently, 100 000 samples were generated, where every 50th draw was stored and the rest discarded to remove the autocorrelation, leaving inference based on 2 000 independent samples. The model was fitted to the data with and without the covariates, to assess whether the covariate effects had any impact on the changing nature of the temporal trend or temporally varying spatial variances in measles susceptibility across Scotland. In common with Section 2.2 each covariate was included in the model as a

natural cubic spline with 3 degrees of freedom, to capture the non-linearity evident in Figure 1.

5.1 Covariate effects

Figure 5 displays the non-linear relationships between each centred covariate and measles susceptibility, with each presented on the odds ratio scale relative to the average value of the covariate. For median house price the odds of susceptibility are at their lowest for houses that are just below the average median house price in Scotland, but as house prices become increasingly cheaper or more expensive the odds of susceptibility increase. For Jobseeker's Allowance the odds of susceptibility are at their lowest in areas with the lowest percentage of the working age population claiming JSA, and increase as the percentage of the population claiming JSA increases. This increase is non-linear, with a linear increase upto a value of the centred covariate just below zero, followed by a flattening off of the slope. The slope then appears to increase again at higher JSA values, but the width of the 95% uncertainty intervals is too wide due to a lack of data to preclude a constant effect. This shape suggests that once the percentage of people in receipt of JSA increases past a certain level the impact on susceptibility of further increasing the percentage diminishes. In terms of the overall effect of socio-economic deprivation on susceptibility, members of the Scottish population claiming JSA would be associated with living in more deprived areas, while areas with a larger median house price are naturally associated with more affluent areas. Hence, the two estimated relationships seem to suggest increases in susceptibility in both very deprived and very affluent areas of Scotland.

18

The proportion of ethnic minority children in each intermediate geography is a proxy for general prevalence of non-white populations, and shows a non-linear relationship with measles susceptibility. The most prominent effect observed in the bottom panel of Figure 5 is that susceptibility increases with increasing proportions of ethnic minorities, and levels out after the centred proportion reaches around 30%. The other interesting feature of this relationship is that a small increased susceptibility is observed for areas that have no ethnic minorities.

Finally, very strong spatial autocorrelation between neighbouring areas even after adjusting for covariates was observed for these data, with a 95% credible interval of $(0.986, 0.994)$ for the spatial autocorrelation parameter ρ, which corroborates the Moran's I results shown in Section 2.
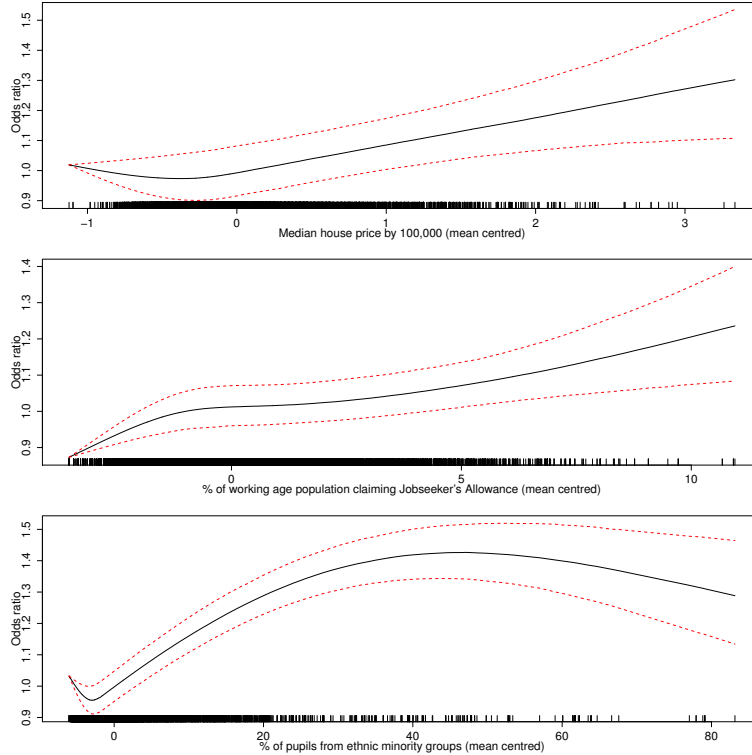
Figure 5: Odds ratio relationships between each covariate and measles suscepti-
bility. In each case the odds ratio is relative to a value of 0 in the mean
centred covariate. The solid lines are the posterior median odds ratios,
while the dashed lines represent 95% credible intervals. The tick marks
along the x-axes indicate data points.

5.2 Overall spatio-temporal patterns

The left half of Figure 6 displays the temporal trend in the average (over
space) estimated proportions of pre-school children susceptible to measles,
which is measured by $\bar{\theta}_t = \frac{\exp(\bar{\mathbf{x}}_t^T \boldsymbol{\beta} + \delta_t)}{1 + \exp(\bar{\mathbf{x}}_t^T \boldsymbol{\beta} + \delta_t)}$, where $\bar{\mathbf{x}}_t = \frac{\sum_{k=1}^{K} \mathbf{x}_{kt}}{K}$. The top
panel displays the results for the model with the covariates, while the
bottom panel relates to a model with no covariates, and both show almost
identical results. An increase in the proportion of pre-school children
susceptible to measles is observed between 1998 and 2004, rising from

under 15% to just over 17.5%. This increase coincides with the period in which the now discredited Wakefield article[1] linking the MMR vaccination with increased risks of autism picked up traction with the Scottish public. The article was partially retracted in 2004, after which we begin to see a decrease in measles susceptibility. By 2014 measles susceptibility in pre-school children is much lower than in 2004, being around 8%.

The right half of Figure 6 displays the temporal trends in the posterior distributions for τ_t^2, which quantifies the amount of spatial variation in measles susceptibility in each year for the model containing the covariates (top panel) and from the model without the covariates (bottom panel). The spatial variability steadily decreases between 1998 and 2006 before leveling off and staying fairly stable between 2006 and 2014. The spatial variability is slightly greater when the covariate effects are not accounted for, with greater spread for most years. Thus there is evidence of a reduction in spatial inequalities in susceptibility between 1998 and 2006, although there appears to be no clear link with the vaccination scare.
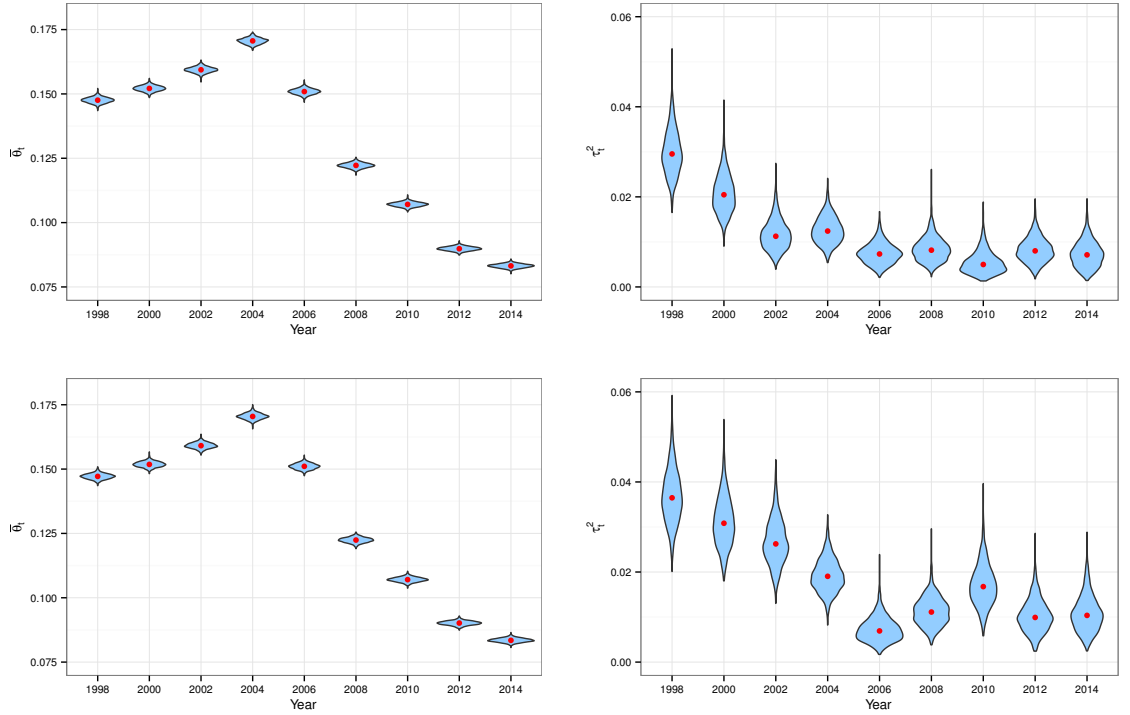
Figure 6: Violin plots from the posterior distributions of the average estimated proportions $\bar{\theta}_t = \frac{\exp(\bar{\mathbf{x}}_t^T \boldsymbol{\beta} + \delta_t)}{1 + \exp(\bar{\mathbf{x}}_t^T \boldsymbol{\beta} + \delta_t)}$ (left column) and the spatial variance (right column) over time. Here $\bar{\mathbf{x}}_t = \frac{\sum_{k=1}^{K} \mathbf{x}_{kt}}{K}$ and the top and bottom panels respectively relate to the model fitted with and without covariates. In all cases the points in each panel are posterior medians.

Spatial maps depicting the mean centred (by year) fitted proportions of children susceptible to measles in Scotland attending pre-school in 1998, 2004, 2008 and 2014 are shown in Figure 7. Measles susceptibility peaks around 2004 – the time the Wakefield article[1] linking the MMR vaccination with autism gained media coverage – with the north west region of Scotland and the remote Shetland islands (far north east) displaying increased susceptibility rates in comparison with the south east. From 2008 onwards measles susceptibility across Scotland has decreased, with susceptibility rates at their lowest in the past 10 years. However as the bottom half of the figure shows, the rural north west of Scotland retains its higher susceptibility rates com-

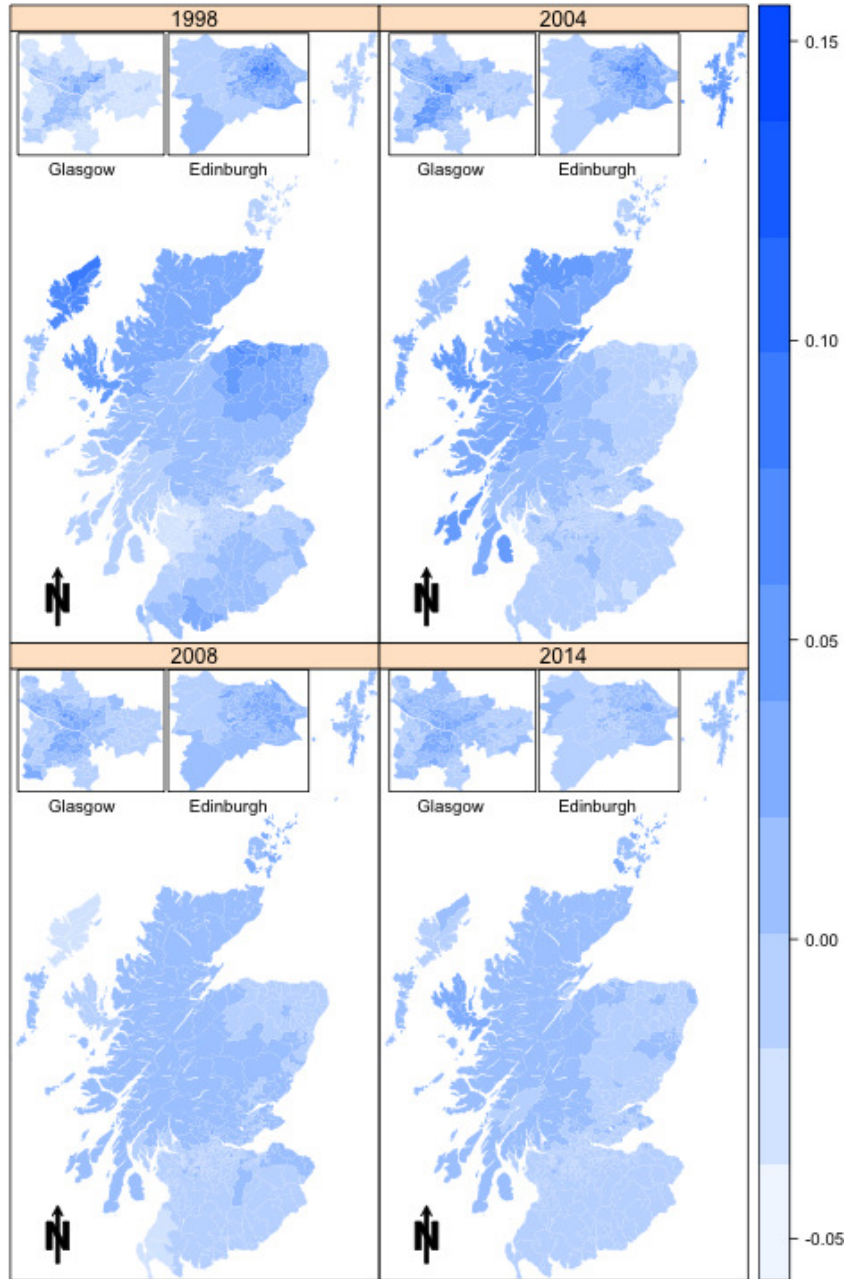pared with the majority of the rest of the country.

Figure 7: Mean centred (by year) fitted proportions of children susceptible to measles in Scotland attending pre-school in 1998, 2004, 2008 and 2014. The panels in the top-left corner highlight the cities of Glasgow and Edinburgh in the centre of Scotland.

23

6   Conclusion

We have presented a spatio-temporal Bayesian hierarchical binomial logistic regression model that allows the magnitude of spatial variability to change over time and the overall temporal trend in disease risk to be estimated. The model was motivated by spatial-temporal data pertaining to the MMR vaccination scare associated with the now retracted Wakefield article[1], which provided an unfounded link between the MMR vaccination and increased risks of autism. The model used in our analysis can be implemented using the `CARBayesST` package, which was developed in conjunction with this paper and is part of the freely accessible statistical programming language `R`. We note that `CARBayesST` can also implement a Poisson log-linear variant of the model.

The results of our study show an increase in measles susceptibility in preschool children between 1998 and 2004, which is concurrent with the time period when the article gained media coverage, thus leading to public fears surrounding the MMR vaccination. The increase in measles susceptibility was most noticeable in the north west region of the country comprised of the highlands and more rural areas, and also the remote Shetland islands in the north east. Our intuition for this pattern would be that general practitioners and health centres are not readily accessible in those areas, so coupled with the negative press associated with the MMR vaccination, parents in those areas would be less likely to get their children vaccinated. Measles susceptibility decreases from 2006 onwards with susceptibility rates at their lowest by 2014 (see Figure 6) at around 8%, which coincides with the Wakefield article's[1] partial retraction in 2004 before being fully discredited in 2010.

24

Our results also show no evidence of an increase in spatial inequality in measles susceptibility between affluent and deprived communities as a result of the autism scare, as the spatial variation consistently decreases between 1998 and 2006, rather than showing any increase. Between 2006 and 2014 the spatial variability remained fairly constant, and was at its lowest level during the entire study period. Three covariates, two used as proxies for measures of deprivation, and one for observing possible ethnic differences were also included in the model. The overall effect of deprivation (as collectively measured by JSA and MHP) on susceptibility appears to be non-linear, with increased susceptibility observed for very poor and very rich areas. The former is expected in common with other national vaccinations (http://www.isdscotland.org/Health-Topics/Child-Health/Publications/2014-03-28/2014-03-28-Immunisation-Report.pdf?46152895690). One possible explanation for the increased susceptibility amongst rich communities is that they be more educated, and hence more aware of the autism link in the media, resulting in reduced vaccination rates. We also observed an increase in susceptibility rates with an increasing percentage of pupils from ethnic minority groups, and a possible explanation for this is that a proportion of these people may be recently arrived immigrants and may not be integrated or familiar with the Scottish vaccination system. There have been no laboratory-confirmed cases of measles in Scotland in 2015 (as of September), with the last laboratory-confirmed case of measles occurring in October 2014. MMR uptake rates for children aged five years of age between 1st April and 30th June 2015 are 97.5% and 93.5% for one dose of MMR (MMR1) and two doses of MMR (MMR2), respectively.

The results in this paper relate to global, that is average across Scotland, trends in measles susceptibility associated with the discredited Wakefield

paper[1], and our future interest is in identifying local intermediate geography level effects. In particular we are interested in identifying which intermediate geographies in Scotland were affected by the scare and which were not, using a spatio-temporal mixture model where the components of the mixture are different shaped temporal trends. For example, one trend shape could be a constant trend over time, while another could be quadratic and exhibit a peak in 2004, and the data in each area would choose the trend component that best fits its temporal susceptibility profile. Another area of future research would be a multivariate space-time analysis for susceptibility among both pre-school and primary school children simultaneously.

Acknowledgements

References

[1] Wakefield A, Murch S, Anthony A, Linnell J, Casson D, Malik M, et al. Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. The Lancet. 1998;351(9103):637–641.

[2] McIntyre P, Leask J. Improving uptake of MMR vaccine. BMJ. 2008;336(7647):729–730.

[3] Pollock K, Potts A, Love J, Steedman N, Donaghy M. Measles in Scotland, 2013. Scottish Medical Journal. 2014;59(1):3–4.

[4] Public Health England. Press Release. National MMR vaccination catch-up programme announced in response to increase in measles cases; 2013. URL: https://www.gov.uk/government/news/national-mmr-vaccination-catch-up-programme-announced-in-response-to-increase-in-measles-cases (accessed 13 July 2015).

[5] Gilberg C, Heijbel H. MMR and autism. Autism. 1998;2(4):423–424.

[6] Taylor B, Miller E, Farrington C, Petropoulos M, Favot-Mayaud I, Li J, et al. Autism and measles, mumps, and rubella vaccine: no epidemiological evidence for a causal association. The Lancet. 1999;353(9169):2026–2029.

[7] Elliman D, Bedford H. MMR: where are we now? Archives of Disease in Childhood. 2007;92(12):1055–1057.

[8] Strebel P, Papania M, Halsey N. Measles vaccine. In: Plotkin SA, Orenstein WA, eds. Vaccines, 4th ed. Philadelphia: WB Saunders; 2004. p. 389–440.

[9] Rue H, Held L. Gaussian Markov Random Fields: Theory and Applications. London: Chapman and Hall-CRC Press; 2005.

[10] Besag J, York J, Mollie A. Bayesian image restoration with two applications in spatial statistics. Annals of the Institute of Statistical Mathematics. 1991;43(1):1–20.

[11] R Core Team. R: A language and environment for statistical computing. Vienna, Austria; 2013. Available from: http://www.R-project.org.

[12] Cleveland W. Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association. 1979;74:829–836.

[13] Moran P. Notes on continuous stochastic phenomena. Biometrika. 1950;37(1/2):17–23.

[14] Ruppert D, Wand M, Carroll R. Semiparametric Regression. Cambridge University Press; 2003.

[15] Woodward W, Gray H, Elliot A. Applied Time Series Analysis. CRC Press; 2011.

[16] Leroux B, Lei X, Breslow N. Estimation of disease rates in small areas: a new mixed model for spatial dependence. In: Statistical models in epidemiology, the environment, and clinical trials. Springer; 1999.

[17] Geweke J. Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. In: IN BAYESIAN STATISTICS. University Press; 1992. p. 169–193.