Published in final edited form as:

Epidemiology. 2017 January ; 28(1): 119–126. doi:10.1097/EDE.0000000000000551.

Spatial variability in the persistence of pneumococcal conjugate vaccine-targeted pneumococcal serotypes among adults

Joshua L. Warren^{1,*}, S. Cassandra Pingali², and Daniel M. Weinberger²

¹Department of Biostatistics, Yale University, New Haven, CT 06520, USA.

²Department of Epidemiology of Microbial Disease, Yale University, New Haven, CT 06520, USA.

Abstract

Invasive pneumococcal disease is a leading cause of morbidity worldwide. Pneumococcal conjugate vaccine effectively reduces the number of cases caused by vaccine-targeted serotypes among children who receive the vaccine and adults who are not directly vaccinated. Recently, there has been debate as to whether adults should receive the same conjugate vaccine as children. In settings where vaccine uptake in children is high, the vaccine serotypes cause a small fraction of disease cases, and direct vaccination might have a small effect. However, direct vaccination might be warranted if geographic regions or subpopulations exist where the targeted serotypes persist at higher levels than expected. To detect such geographic variability, new methodology is required. We introduce an innovative, spatially varying change points model, combined with spatially varying intercepts and slopes, to jointly determine whether the beginning date of the vaccine-associated decline, the initial baseline proportion of invasive pneumococcal disease cases caused by vaccine-targeted serotypes, and/or the rate of decline of vaccine-targeted serotypes vary in the adult population across Connecticut, 1998-2009. Results indicate that there is substantial spatial variability in the pattern with which vaccine-targeted serotypes decline, suggesting that the fraction of invasive pneumococcal disease cases that could have been preventable by direct vaccination of adults in Connecticut during the study period differed over time and space. The newly developed model is shown to outperform a number of competitors in terms of explanatory and predictive ability.

Introduction

Pneumococcal disease, a common infection caused by the bacterium *Streptococcus pneumoniae*, is a leading cause of morbidity across the world. A subset of the over 90 known pneumococcal serotypes cause the majority of invasive pneumococcal disease cases, a severe form of the disease which is defined by the isolation of bacteria from the blood or cerebrospinal fluid. Pneumococcal conjugate vaccine is associated with reduced rates of invasive pneumococcal disease caused by the serotypes targeted by the vaccine. This reduction was observed among children who received the vaccine and within the adult population, which was not vaccinated but which benefitted indirectly.¹⁻⁵

*To whom correspondence should be addressed: phone: 203-785-4188; joshua.warren@yale.edu.

The Advisory Committee on Immunization Practices recently recommended the use of pneumococcal conjugate vaccines for all adults over the age of 65 and for those aged 19-64 with certain risk factors.⁶⁻⁷ Direct vaccination of adults generally is associated with a diminishing impact over time as the indirect protection that results from vaccinating children increases; however, there may be subsets of the adult population or geographic regions where this indirect benefit is less pronounced and, therefore, where direct vaccination is more important. One factor that could influence the strength of indirect benefit received by adults is the rate of vaccine uptake among children in surrounding areas. Low vaccine uptake could allow the vaccine-targeted serotypes to persist among adults. Factors associated with the socioeconomic status (SES) of a region may also influence the serotype distribution and transmission patterns in localized populations.

In this paper, we explore the spatial variability of invasive pneumococcal disease cases caused by serotypes targeted by the original conjugate vaccine (PCV7) among adults aged 18+ in Connecticut, 1998-2009. Our goal is to determine if certain geographic regions have (i) higher than expected baseline proportions of invasive pneumococcal disease cases caused by vaccine-targeted serotypes in the pre-vaccine period, (ii) slower rates of decline of these proportions in the post-vaccine period, and/or (iii) later than expected dates when this decline begins, after adjusting for vaccine uptake and SES characteristics. Fully characterizing the spatial variability and uncertainty in these data will help to determine if there are areas where the vaccine-targeted serotypes persist at a higher level or where pediatric vaccination rates are sub-optimal to protect adults indirectly.

We introduce a newly developed model that can simultaneously address each of the proposed questions within a unified Bayesian framework. Our individual-level probit change points regression model includes spatially varying intercepts and slopes in order to explore the spatial variability in baseline proportions and rates of decline, respectively. Spatially varying change points are introduced in order to account for geographic differences in the date of beginning decline of the proportions. To our knowledge, this is the first use of spatially correlated change points in the literature. The introduced model is applied to invasive pneumococcal disease data from Connecticut in 1998-2009 and is shown to outperform competing models in terms of explanatory ability and predictive performance.

Methods

Data Sources

Data for this study are available at multiple geographic scales including the individual-level, census tract, United States (US) Postal Service ZIP code service area, and ZIP Code Tabulation Area (ZCTA). ZCTAs are areas that the US Census Bureau defines based on the ZIP code service areas and in most cases, the ZCTA code matches the ZIP code for an area.⁸ We geographically link these different sources of information where possible.

Data on individual-level adult invasive pneumococcal disease cases were collected from the Active Bacterial Core surveillance system in Connecticut from 1998-2009. For each individual diagnosed with invasive pneumococcal disease, information was collected on their age group (18-39, 40-64, 65+), month and year that the sample was collected, the

pneumococcal serotype that caused the case of invasive pneumococcal disease, and the census tract and ZIP code of residence at the time of diagnosis (n=3,836).

We estimated ZIP code-level PCV7 uptake information among children using data from the Connecticut Immunization Registry and Tracking System. For each ZIP code during each month/year, we estimated the number of children eligible to receive PCV7 and the number of children who received three and/or four doses using a cohort approach while assuming a constant birth rate over the year and that vaccinations for children who received three or four doses were administered according to schedule.^{5,9}

SES indicators were collected from the US Census Bureau at the tract-level based on the 2000 census. Considered covariates include population density, percent of the tract that is black, Hispanic, and white, income per capita, and total population size. After initial exploratory modeling of the data, only population density was included in the final set of covariates.

In the final analysis, the following covariates are included: age category of the individual (18-39, 40-64, and 65+); log of the tract-level population density (standardized); ZIP code-level information on vaccine uptake among children. We include the proportion of children who received at least three doses of PCV7 along with the proportion of children who received three but not four doses of PCV7. Both of these variables are standardized at each month/year by subtracting off the respective proportion averaged across all ZIP codes at that date. Including these variables allows us to investigate the benefit of giving children three versus four doses of the vaccine with respect to protection of the adult population.

We note that the population density and vaccine uptake variables are not available at the individual-level and people living in the same geographic region will have the same values of these covariates. Therefore, the results are interpreted at the geographic region-level for these covariates. Individuals with missing information on the final set of included covariates are removed from the analysis (n=315) resulting in a total sample size of 3,521. This includes individuals from 236 out of the total 266 ZCTAs in Connecticut.

eFigure 1 displays the number of invasive pneumococcal disease cases in each ZCTA for the final analysis dataset while Figure 1 shows the proportion of invasive pneumococcal disease cases caused by a targeted serotype at each year of the study, averaged across all ZCTAs. Table 1 reports the characteristics of the study population. The study was approved by the Human Investigation Committees at Yale and the Connecticut Department of Public Health (DPH). Certain data used in this publication were obtained from DPH. The authors assume full responsibility for analyses and interpretation of these data.

Statistical Model

We introduce a newly developed model for analyzing the persistence of pneumococcal conjugate vaccine-targeted pneumococcal serotypes among adults. In the model, each geographic region has a unique intercept, slope, and change point parameter. Estimating these parameters enables the identification of regions where the proportion of invasive pneumococcal disease cases caused by vaccine-targeted serotypes persists longer than

expected (large change point, slope closer to zero) or that have higher baseline probabilities than expected (large intercept). The addition of these spatially varying parameters increases the flexibility of the model while not imposing any unwarranted assumptions on its structure. It also allows us to simultaneously investigate each scientific question of interest regarding the persistence of vaccine-targeted serotypes.

At a date occurring before the location-specific change point, the probability that an individual's case of invasive pneumococcal disease is caused by a targeted serotype is modeled as a function of person/location/time-specific covariates (e.g. pneumococcal conjugate vaccine uptake, population density) and spatial intercept alone. This represents the period of time before the vaccine has had any impact on the adult population within a specific region. Including location-specific intercepts allows for the possibility that the baseline probability of a case of invasive pneumococcal disease being caused by a targeted serotype varies by region. Including location-specific change points allows for the date of beginning decline of the rates to change across the regions if necessary.

At a date occurring after the location-specific change point, the probability that an individual's case of invasive pneumococcal disease is caused by a targeted serotype is modeled as a function of person/location/time-specific covariates, spatial intercept, and a location-specific linear time trend. This represents the period of time after the vaccine is introduced to children when the indirect benefits begin to impact adults. The location-specific slope parameters allow for the possibility that the rate of decline is changing across the regions.

The model is formally defined as $Y_i\{z(i), t(i)\} | p_i\{z(i), t(i)\} \sim \text{Bernoulli}[p_i\{z(i), t(i)\}]$, $i = 1, \dots, n$, where $Y_i\{z(i), t(i)\}$ represents the response from individual i located in ZCTA $z(i)$ at month/year $t(i)$ of the study and is equal to one if the invasive pneumococcal disease case is caused by a targeted serotype and zero otherwise. Multiple individuals in the study are often located within the same ZCTA such that $z(i) \in \{z_1^*, \dots, z_m^*\}$ for all i where z_j^* represents a unique ZCTA, m is the number of unique regions observed in the study ($m=236$), n is the total sample size ($n=3,521$), and generally $m \ll n$. We observe information on invasive pneumococcal disease cases at discrete monthly time periods beginning with January 1998 ($t=1$) and extending through December 2009 ($t=144$). Multiple individuals in the study are observed at the same monthly time period such that $t(i) \in \{1, \dots, 144\}$ for all i . The $z(i)$ and $t(i)$ functions map an individual to the correct ZCTA and monthly time period, respectively.

We model the underlying probability that the invasive pneumococcal disease case for individual i is caused by a vaccine-targeted serotype using a newly developed probit change points regression framework such that

$$\Phi^{-1}[p_i\{z(i), t(i)\}] \quad (1)$$

$$= \begin{cases} \mathbf{x}_i^T \boldsymbol{\gamma} + \tilde{\beta}_0\{z(i)\} & , t(i) \leq \theta\{z(i)\}; \\ \mathbf{x}_i^T \boldsymbol{\gamma} + \tilde{\beta}_0\{z(i)\} + \tilde{\beta}_1\{z(i)\}[t(i) - \theta\{z(i)\}] & , t(i) \geq \theta\{z(i)\}. \end{cases}$$

ZCTA-specific intercepts, slopes, and change points are denoted by $\tilde{\beta}_0\{z(i)\}$, $\tilde{\beta}_1\{z(i)\}$, and $\theta\{z(i)\}$, respectively. We center time at zero using the change point ($t(i) - \theta\{z(i)\}$) in order to connect the piecewise components at the change point and create a continuous process across time. The vector of covariates specific to individual i at location $z(i)$ and time $t(i)$ is given by $\mathbf{x}_i$ and $\boldsymbol{\gamma}$ represents the associated vector of regression parameters.

Prior Specification

We allow for the possibility that spatially referenced parameters associated with neighboring ZCTAs are more similar through use of a spatially structured prior distribution. We also believe that there may be correlation between the change point, intercept, and slope specific to a single ZCTA. Therefore, we utilize a multivariate spatial prior distribution to incorporate this potential source of correlation.

PCV7 was first introduced in 2000 (Figure 1), therefore, a change point cannot occur before January 2000 ($t=25$). In order to observe the change point during the time frame of our study, it would have to occur on or before November 2009 ($t=143$). A change point occurring in the final month would not be detected given the data we have access to. Therefore, we assume that each ZCTA-specific change point is a continuous valued parameter in that range such that $\theta(z_j^*) \in [25, 143]$ for all j . In order to introduce the possibility of spatial correlation between these parameters, we transform them to the real

line such that $\tilde{\psi}(z_j^*) = \ln \left(\frac{\theta(z_j^*) - 25}{143 - \theta(z_j^*)} \right) \in \mathbb{R}$ and model the transformed parameters using a

constant overall mean and spatial deviation such that $\tilde{\psi}(z_j^*) = \psi + \psi(z_j^*)$. The spatially varying intercepts and slopes are similarly decomposed into two pieces representing the overall mean of the process and location-specific deviation such that $\tilde{\beta}_k(z_j^*) = \beta_k + \beta_k(z_j^*)$, $k = 0, 1$ for all j .

In order to account for potential correlation between the change point, intercept, and slope specific to a ZCTA as well as spatial correlation between the ZCTAs, we specify a multivariate conditional autoregressive model for the vector of spatial deviations,

$\boldsymbol{\lambda} = \left\{ \lambda(z_1^*)^T, \dots, \lambda(z_m^*)^T \right\}^T$, where $\lambda(z_j^*) = \left\{ \psi(z_j^*), \beta_0(z_j^*), \beta_1(z_j^*) \right\}^T$.¹⁰ This distribution can be better understood conditionally such that

$$\lambda(z_j^*) | \lambda_{-j} \sim \text{MVN}\left(\frac{1}{w_{j+} + k} \sum_{k=1}^m w_{jk} \lambda(z_k^*), \frac{1}{w_{j+}} \Sigma_\lambda\right)$$

where λ_{-j} is λ with $\lambda(z_j^*)$ removed, $\text{MVN}(\mu, \Sigma)$ represents the multivariate normal distribution with mean vector μ and covariance matrix Σ , w_{jk} describes the neighborhood association between ZCTA j and ZCTA k , w_{j+} represents the sum of the neighborhood weights for ZCTA j , and Σ_λ is a three by three covariance matrix for the change point, intercept, and slope.

We define two ZCTAs to be neighbors if they share a common border. Therefore, $w_{jk} = 1$ if ZCTA j and k share a common border and $w_{jk} = 0$ otherwise. By definition, $w_{jj} = 0$ for all j . This prior distribution allows for neighboring regions to share information for estimation if appropriate and for the possibility of cross-covariance between the parameters. Σ_λ controls the level of spatial information sharing as well as the cross-covariance between the spatially referenced parameters.

In order to complete the model specification, we assign prior distributions to the remaining parameters. The covariate regression parameters (γ_j) and the non-spatial change point, intercept, and slope parameters (ψ, β_0, β_1) are given independent and identically distributed normal distributions with large, fixed prior variance such that $\gamma_j, \psi, \beta_0, \beta_1 \sim N(0, 10^3)$ for all j . This represents a vague yet proper prior distribution and reflects our lack of knowledge regarding the values of those parameters. A Wishart prior distribution with scale matrix equal to the three by three identity matrix and degrees of freedom equal to four is assigned to Σ_λ^{-1} . This vague yet proper prior distribution allows for marginally uniform prior cross-correlations between the spatial change point, intercept, and slope parameters.¹¹

Sensitivity Analysis

We assess the sensitivity of our results to the choice of a fixed, large variance for the non-spatial change point, intercept, and slope parameters. We specify $\psi, \beta_0, \beta_1 \sim N(0, \tau^2)$ where τ^2 is given a weakly informative inverse gamma prior distribution such that $\tau^2 \sim \text{IG}(0.001, 0.001)$. We then compare these results to the fixed variance results using model comparison metrics and plots of estimated posterior means.

Computing

All models are fit using R statistical software (R Development Core Team, Vienna, Austria).¹² Posterior inference is based on 85,000 posterior samples from three chains after a burn-in period of 15,000 iterations in each chain for a total of 255,000 posterior samples. We assessed convergence through visual inspection of the trace plots and calculation of the potential scale reduction factor for each model parameter¹³⁻¹⁴. The average, median, minimum, and maximum values of the potential scale reduction factor across all model parameters was 1.01, 1.00, 1.00, and 1.08 respectively. Potential scale reduction factor values near one indicate that the chains have likely converged to their target distribution. The

number of posterior samples to collect after burn-in was determined based on the analysis of batch mean Monte Carlo standard errors¹⁵. Model fitting details can be seen in eSection 1.

Summarizing Output

We develop a new metric to analyze the relative ranking of each ZCTA with respect to values of the change points, intercepts, and slopes jointly. For each collected posterior sample of model parameters, each ZCTA is given three rankings based on its change point, intercept, and slope parameters. We subtract one from each ranking, divide each by $m - 1$, and average these scaled rankings for each ZCTA. ZCTAs with average values closer to one tend to have a combination of spatial parameters that are extreme (large change point, large intercept, slope closer to zero) and may be important regions to focus on when investigating adult vaccination.

We formally compare the performance of competing models using the deviance information criterion and a posterior predictive loss criterion (D_k).¹⁶⁻¹⁷ Deviance information criterion is a common model comparison tool used in the Bayesian setting that balances model fit ($\bar{D}$: posterior mean of the deviance statistic) with a measure of complexity (p_D : effective number of parameters) where $DIC = \bar{D} + p_D$ and smaller values are preferred. D_k is a model comparison tool based on posterior predictive loss that also attempts to balance model fit with complexity. In our setting, we rely on the deviance of the Bernoulli distribution and set $k = 10^{10}$ to calculate D_k . As with deviance information criterion, smaller values of D_k are preferred. Deviance information criterion is useful when the explanatory ability of a model is the main interest of an analysis while D_k is often preferred if the predictive ability of a model is the focus.

Simulation Study

We conduct a simulation study to assess the performance of the spatially varying change points model. Full details are presented in eSection 2.

Results

Vaccine-targeted IPD Rates in CT, 1998-2009

In Table 2, we display the posterior inference for the non-spatial model (1) parameters. In eTable 2, we present the posterior summaries for entries of Σ_λ . None of the included SES covariates are associated with the probability that an invasive pneumococcal disease case is caused by a vaccine-targeted serotype. The slope after the change point is negative, indicating that the proportion of cases caused by vaccine-targeted serotypes decreased after the change point (Figure 1). As the proportion of children in a ZCTA who received three but not four doses of the vaccine increases, the probability that an invasive pneumococcal disease case in an adult is caused by a targeted serotype is also increased. This suggests that completing the vaccination cycle by receiving the fourth (booster) dose of the vaccine may provide an indirect benefit to the adult population.

Figure 2 displays the posterior mean estimates for the change point $\left(\theta\left(z_j^*\right)\right)$, intercept $\left(\tilde{\beta}_0\left(z_j^*\right)\right)$, and slope $\left(\tilde{\beta}_1\left(z_j^*\right)\right)$ parameters. The corresponding posterior standard deviation estimates are presented in eFigure 2. Figure 2B suggests that in the southwestern region of Connecticut, the initial proportion of invasive pneumococcal disease cases caused by vaccine-targeted serotypes is higher than in other parts the state. Considering Figure 2A and 2B jointly indicates that generally, in the areas of the state where the intercepts are largest, the change points tend to occur earlier. Note that the change points across the state are centered around 39 months, which agrees with the period of decline in Figure 1. The rates of decline in vaccine-targeted invasive pneumococcal disease cases appear to be slower in the eastern part of Connecticut. Figure 2D displays the posterior means of the relative ranking metric and allows us to jointly identify regions with an extreme set of spatial parameters indicating increased risk (values closer to one).

In Figure 3, we display the predicted probability that an invasive pneumococcal disease case is caused by a targeted serotype for an adult aged 40-64 living in an area with average population density for a number of different time points. All predictions account for vaccine uptake information. The corresponding posterior standard deviation estimates are presented in eFigure 3. In early 1998, a majority of the high probability areas are seen in the southwest region of Connecticut, mainly due to the high pre-vaccine intercepts. By late 2009, the high probability regions are seen in the eastern part of the state due to the late-occurring change points and slower rates of decline. The spatial pattern of predictions is better understood when Figures 2 and 3 are analyzed together.

Model Comparisons

We compare the newly developed model with a number of alternatives with respect to model fit and predictive ability. We first explore the most appropriate temporal model for the data by comparing a standard change point model, a linear time trend model, and a model that flexibly incorporates time using a cubic spline with five degrees of freedom. Next, we consider the original framework in (1) while allowing for all combinations of spatially varying change points, intercepts, and/or slopes.

In Table 3, we display all model comparison results. Focusing first on understanding the temporal structure of the data, the results suggest that the standard change point model provides improved explanatory and predictive ability over the linear time trend and cubic spline models. As expected, the linear time trend model performs poorly because it ignores the pre-vaccine period of stable invasive pneumococcal disease rates seen in Figure 1. The cubic spline model provides an improvement in fit and prediction but uses an increasing number of effective parameters (p_D) to do so.

Next, we attempt to determine if any of the parameters are constant across space. The findings suggest that the full model in (1) represents the best overall model for explanatory and predictive purposes and that allowing each component to vary spatially represents an improvement over competing models. The models that include spatially varying change points always outperform the respective models with a constant change point.

To further explore the impact of ignoring spatial variability in the data, we compare the predictive performance of the competing non-spatial models with the full model in (1). For each individual in the study, we obtain samples from the posterior distribution of $p_i\{z(i), \ell(i)\}$ and summarize them using the posterior mean. We average these posterior means across space and time, and compare them to the observed data averaged in the same way. In eFigure 4, we display the yearly comparisons averaged across all ZCTAs. Each point in the figure represents a year in the study. Other than the linear time trend model, each model provides similar and adequate predictive ability across time. In eFigure 5, we display the ZCTA comparisons averaged across time. Each point in the figure represents a ZCTA averaged across all time points. Each of the non-spatial models provides similar predictions and fails to adequately handle the spatial variability. The correlation between observed and predicted proportions range from 0.34 to 0.37 for these models. The full model in (1) however, is able to substantially improve prediction and results in a correlation of 0.83.

Finally, we compare how the predictions vary due to unexplained spatial variability and due to differences in vaccine uptake between ZCTAs. eFigure 6 displays these results and fully describes the findings. The results indicate that a majority of the variability in predictions between ZCTAs is due to unexplained spatial variability.

Sensitivity Analysis

Deviance information criterion, p_D , and D_k from the sensitivity analysis model are 4085.6, 61.3, and 3016.1 respectively, compared with 4086.2, 61.0, and 3014.9 from the final model (Table 3). The posterior means from all model parameters are also in close agreement between the two models as shown in eFigure 7. The results do not appear to be sensitive to this choice of prior.

Simulation Study

The simulation study results suggest that Method 1 consistently provides adequate coverage of the change points under all considered designs and, with any form of spatial correlation between the change point parameters, provides an improved fit to the data. A full discussion of the results is provided in eSection 2.

Discussion

The statistical literature on change point models is vast and a number of review articles on the topic are available.¹⁸⁻²⁰ However, relatively few studies have addressed change point modeling for spatially referenced data. The majority of previously introduced change point regression models assume that a single change point is shared across all responses. This assumption may be inappropriate if the observations are arising from different spatial locations. In this work, we relaxed this assumption by allowing for spatially correlated location-specific change points.

The use of Bayesian change point models while accounting for potential spatial correlation between responses has been considered in past studies where the authors included spatially and spatiotemporally correlated adjustments to the intercept.²¹⁻²³ In another study, the authors increased flexibility of the change point framework by allowing for different

independent spatiotemporal adjustments to the intercept before and after the change point occurred.²⁴ In a more recent study, the authors introduced a model with spatially varying intercepts, slopes, and change points.²⁵ The prior distribution for these parameters allowed for correlation among parameters at a single spatial location but assumed independence between the parameters from different locations.

We found that substantial spatial variability remained even after adjusting for a number of spatially varying covariates. This indicates that additional, unexplained spatial factors might influence the serotype distribution and that an individual's risk of infection depends not just on their ZCTA of residence but also on their broader region of residence. Commuting or other regional interaction patterns could explain some of the variability. Likewise, the spatial variation could indicate the presence of unmeasured spatial covariates, such as air pollution. Future work could focus on applying the model in new data settings as well as extending the methodology to account for multiple change points for a single response, spatiotemporal change points for cyclical processes, and/or multivariate change point models when working with multivariate outcomes.

Using our model, we were able to show that substantial variability exists in the proportion of invasive pneumococcal disease cases caused by vaccine-targeted serotypes across time and space. Results suggest that the date of beginning decline, the rate of decline, and the initial baseline proportion of vaccine-targeted cases varied across Connecticut. Vaccine uptake among children was shown to be the main predictor of vaccine-targeted invasive pneumococcal disease proportions in the adult population while four doses may be more important for indirect protection of the adult population. Results suggest that direct vaccination of the adult population in Connecticut during the study period would have been more impactful in the geographic areas with a combination of large baseline proportions, late change points, slow rates of decline, and low vaccine uptake.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Nancy Sharova and Susan Petit (CT Department of Public Health) for providing access to the disease and vaccination data and for discussions of issues related to the data.

Conflicts of Interest and Financial Support: This work was supported by an investigator-initiated research grant from Pfizer to Yale University (WI 178463). The funder had no role in the design or conduct of the study or the decision to publish. The work was also partially funded by: #P30AG021342 NIH/NIA (Scholar at the Claude D. Pepper Older Americans Independence Center at Yale University School of Medicine), UL1 TR000142, The Bill and Melinda Gates Foundation (OPP1114733), and R56 AI110449-01A1.

References

1. Lexau CA, Lynfield R, Danila R, et al. Changing epidemiology of invasive pneumococcal disease among older adults in the era of pediatric pneumococcal conjugate vaccine. *JAMA*. 2005; 294:2043–2051. [PubMed: 16249418]

2. Ardanuy C, Tubau F, Pallares R, et al. Epidemiology of invasive pneumococcal disease among adult patients in Barcelona before and after pediatric 7-valent pneumococcal conjugate vaccine introduction. 1997-2007. *Clinical Infectious Diseases*. 2009; 48:57–64. [PubMed: 19035779]
3. Pilishvili T, Lexau C, Farley MM, et al. Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. *Journal of Infectious Diseases*. 2010; 201:32–41. [PubMed: 19947881]
4. Feikin DR, Kagucia EW, Loo JD, et al. Serotype-specific changes in invasive pneumococcal disease after pneumococcal conjugate vaccine introduction: a pooled analysis of multiple surveillance sites. *PLoS Medicine*. 2013; 10:e1001517. 2013. [PubMed: 24086113]
5. Pingali SC, Warren JL, Mead AM, Sharova N, Petit S, Weinberger DM. Association between local pediatric vaccination rates and patterns of pneumococcal disease in adults. *The Journal of Infectious Diseases*. 2015:jiv431v1–jiv431.
6. Centers for Disease Control and Prevention. Use of 13-valent pneumococcal conjugate vaccine and 23-valent pneumococcal polysaccharide vaccine for adults with immunocompromising conditions: recommendations of the Advisory Committee on Immunization Practices (ACIP). *Morbidity and Mortality Weekly Report*. 2012; 61:816–819. [PubMed: 23051612]
7. National Foundation for Infectious Diseases. [April 13, 2015] Pneumococcal disease. <http://www.nfid.org/idinfo/pneumococcal>.
8. United States Census Bureau. [May 5, 2015] ZIP Code Tabulation Areas (ZCTAs). <https://www.census.gov/geo/reference/zctas.html>.
9. Nuorti JP, Martin SW, Smith PJ, Moran JS, Schwartz B. Uptake of pneumococcal conjugate vaccine among children in the 1998–2002 United States birth cohorts. *American Journal of Preventive Medicine*. 2008; 34:46–53. [PubMed: 18083450]
10. Mardia KV. Multi-dimensional multivariate Gaussian Markov random fields with application to image processing. *Journal of Multivariate Analysis*. 1988; 24:265–284.
11. Gelman, A., Carlin, JB., Stern, HS., Dunson, DB., Vehtari, A., Rubin, DB. *Bayesian Data Analysis*. Chapman & Hall/CRC texts in statistical science; London: 2013.
12. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. 2014. <http://www.R-project.org/>
13. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Statistical Science*. 1992; 7:457–472.
14. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*. 1998; 7:434–455.
15. Roberts, GO. Markov chain concepts related to sampling algorithms.. In: Gilks, W.Richardson, S., Spiegelhalter, D., editors. *Markov Chain Monte Carlo in Practice*. Chapman & Hall; London: 1996. p. 45-58.
16. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2002; 64:583–639.
17. Gelfand AE, Ghosh SK. Model choice: a minimum posterior predictive loss approach. *Biometrika*. 1998; 85:1–11.
18. Zacks S. Survey of classical and Bayesian approaches to the change-point problem: fixed sample and sequential procedures of testing and estimation. *Recent Advances in Statistics*. 1983; 25:245–269.
19. Wolfe DA, Schechtman E. Nonparametric statistical procedures for the changepoint problem. *Journal of Statistical Planning and Inference*. 1984; 9:389–396.
20. Siegmund D. Boundary crossing probabilities and statistical applications. *The Annals of Statistics*. 1986; 14:361–404.
21. Beckage B, Joseph L, Belisle P, Wolfson DB, Platt WJ. Bayesian change-point analyses in ecology. *New Phytologist*. 2007; 174:456–467. [PubMed: 17388908]
22. Cai B, Lawson AB, McDermott S, Aelion CM. A Bayesian semiparametric approach with change points for spatial ordinal data. *Statistical Methods in Medical Research*. 2012; 0:1–15.
23. Yu Q, Scribner R, Carlin B, et al. Multilevel spatio-temporal dual changepoint models for relating alcohol outlet destruction and changes in neighbourhood rates of assaultive violence. *Geospatial Health*. 2008; 2:161–172. [PubMed: 18686265]

24. Majumdar A, Gelfand AE, Banerjee S. Spatio-temporal change-point modeling. *Journal of Statistical Planning and Inference*. 2005; 130:149–166.
25. Wagner T, Midway SR. Modeling spatially varying landscape change points in species occurrence thresholds. *Ecosphere*. 2014; 5:1–16.

Figure 1. Adult invasive pneumococcal disease proportions over time

Proportion of adult invasive pneumococcal disease cases caused by a vaccine-targeted serotype in each year of the study, averaged across all Connecticut ZIP Code Tabulation Areas. The dashed line indicates the year that PCV7 was introduced to children (2000).

Figure 2. Posterior means of spatial parameters

Posterior mean estimates from the model in (1) of the spatially varying (A) change points (months since December 1997), (B) intercepts, (C) slopes, and (D) combined ranking metric.

Figure 3. Model predictions over time

Posterior mean predictions from the model in (1) of the probability that an invasive pneumococcal disease case is caused by a vaccine-targeted serotype (adult aged 40-64 in a census tract with average population density) for (A) January 1998, (B) March 2001, (C) January 2005, and (D) December 2009.

Table 1

Study population characteristics. IPD: invasive pneumococcal disease.

Characteristic	IPD Case Serotype	
	Vaccine-Targeted	Other
Sample Size, No.	1,686	1,835
Age Category, No.:		
18-39	204	176
40-64	601	736
65+	881	923
Race/Ethnicity %^a:		
White	77.9	77.5
Black	13.3	13.8
Hispanic	7.0	6.6
Population Density^b	27.4	27.3
Income Per Capita^b	10.1	10.0

^aPercent of tract.^bLog-scale, tract-level

Table 2

Posterior summaries for the non-spatial model parameters. SD: posterior standard deviation.

Parameter	Quantiles			
	Mean	SD	0.025	0.975
Intercept	0.50	0.08	0.35	0.65
Population Density^a	-0.00	0.03	-0.06	0.06
Age Group:				
40-64	-0.10	0.08	-0.26	0.05
65+	-0.04	0.08	-0.19	0.12
Vaccine Uptake %:				
3 Doses^b	-0.31	0.78	-1.8	1.2
3, Not 4 Doses^b	2.2	0.94	0.31	4.0
Post Change Point Slope	-2.1	0.17	-2.5	-1.8

^aLog-scale; standardized by subtracting off the mean and dividing by the standard deviation of the variable.

^bStandardized at each month by subtracting off the respective Connecticut average.

Table 3

Model comparison results. The deviance information criterion (DIC), effective number of parameters (p_D), and posterior predictive loss criterion (D_k) are presented with smaller values of DIC and D_k being preferred. CP: change point.

Model	DIC(p_D)	D_k
Linear Time Trend	4175.4 (7.0)	3138.1
Cubic Spline, 5 df	4124.3 (11.0)	3094.3
Standard CP	4118.8 (8.1)	3092.6
Spatial CPs	4115.4 (22.5)	3078.2
Standard CP, Spatial Slopes	4114.9 (26.9)	3077.3
Standard CP, Spatial Intercepts	4101.9 (37.7)	3050.6
Standard CP, Spatial Coefficients	4099.6 (67.7)	3018.7
Spatial CPs, Spatial Slopes	4100.2 (29.0)	3062.9
Spatial CPs, Spatial Intercepts	4098.9 (57.8)	3026.6
Spatial CPs, Spatial Coefficients	4086.2 (61.0)	3014.9