

MULTIVARIATE SPATIOTEMPORAL MODELING OF AGE-SPECIFIC STROKE MORTALITY

BY HARRISON QUICK, LANCE A. WALLER AND MICHELE CASPER

*Drexel University, Emory University and the Centers
for Disease Control and Prevention*

Geographic patterns in stroke mortality have been studied as far back as the 1960s when a region of the southeastern United States became known as the “stroke belt” due to its unusually high rates. While stroke mortality rates are known to increase exponentially with age, an investigation of spatiotemporal trends by age group at the county level is daunting due to the preponderance of small population sizes and/or few stroke events by age group. In this paper, we implement a multivariate space–time conditional autoregressive model to investigate age-specific trends in county-level stroke mortality rates from 1973 to 2013. In addition to reinforcing existing claims in the literature, this work reveals that geographic disparities in the reduction of stroke mortality rates vary by age. More importantly, this work indicates that the geographic disparity between the “stroke belt” and the rest of the nation is not only persisting, but may in fact be worsening.

1. Introduction. Stroke (i.e., cerebrovascular disease) is the fourth leading cause of death in the United States (US) and third—behind heart disease and cancer—among those aged 85 and older, with rates increasing exponentially with age [Xu et al. (2016)]. Geographic patterns of stroke mortality have been studied as far back as the 1960s when Borhani (1965) identified a region of the southeastern US stretching from Mississippi to North Carolina which had the highest rates of stroke mortality—a region which would become known as the “stroke belt.” Later work by Casper et al. (1995) noticed an apparent shift in the stroke belt, observing that parts of the Mississippi River Valley appeared in the highest decile of mortality rates in the early 1990s where they had previously not. More recently, Schieb et al. (2013) studied geographic trends in stroke hospitalizations from 1995 to 2006 and noted that this shift in the stroke belt had persisted among Medicare beneficiaries ages 65 and older, stretching further into parts of Texas and Oklahoma. In addition to changing geographic patterns, numerous studies have observed the overall declines in stroke mortality [e.g., Gillum, Kwagyan and Obiesesan (2011), Howard et al. (2001)]. While many of these studies have age-adjusted their data—indirectly accounting for disparities in stroke mortality across age by accounting

Received August 2016; revised June 2017.

Key words and phrases. Age disparities in health, Bayesian methods, geographic disparities in health, nonseparable models, small area analysis.

for the counties' age distributions—this has precluded inference within individual age groups.

There are multiple potential motivations for this aggregation and data standardization step. For instance, the issue of small population sizes and/or low counts in many US counties can lead to unreliable estimates of stroke mortality rates. This (along with computational burden) is only exacerbated when the data are stratified by a factor such as age group. Rather than aggregate data, however, we will investigate spatiotemporal trends in stroke mortality by jointly modeling data from three age-based subpopulations, permitting inference at the county level for each age group while preserving the ability to compute age-adjusted rates. In addition to the challenges of small area estimation, the decision to use age-aggregated data may also be due to the availability of public-use data. While we will revisit this issue in Section 5, we will assume from this point forward that access to these data is not an issue.

To analyze multivariate spatial data, it is common to look toward variations of the conditional autoregressive (CAR) model [Besag (1974), Besag, York and Mollié (1991)] or its multivariate generalization—the multivariate CAR (MCAR) of Gelfand and Vounatsou (2003). In this paper, we will implement the multivariate space–time CAR (MSTCAR) model of Quick, Waller and Casper (2017a) to investigate age-specific spatiotemporal trends in the stroke mortality data described in Section 2. Details regarding the MSTCAR model and its extension to the generalized linear model setting to analyze rare event count data are provided in Section 3. In particular, the MSTCAR model allows for a *nonseparable* model structure which permits a between-age covariance structure which evolves over time and age-specific temporal correlation parameters. This is in contrast to *separable* models which prohibit temporal evolution in the between-age covariance structure and require identical temporal correlation parameters among the various age groups. Other nonseparable methods [e.g., Martinez-Beneito (2013)] allow for varying spatial structures by utilizing *proper* MCAR models; while such approaches are feasible when the dimension of the spatial domain is small, the large number of counties in the US prohibits the use of proper MCAR models. Therefore, our approach strikes a necessary balance between computational burden and model flexibility to provide a more accurate portrayal of the spatiotemporal trends in stroke mortality by age group. Our analysis of the stroke mortality data is presented in Section 4, where we observe spatiotemporal trends which vary by age group. In particular, we find that rates for those aged 65–74 exhibit a stronger degree of spatial clustering with larger geographic disparities than rates for those 75 and older. Finally, we summarize our findings and offer some concluding remarks in Section 5.

2. Data description. The study population for this analysis includes all US residents aged 65 or older. In order to assess differences across the high-risk age ranges, the data were separated into $N_g = 3$ groups: those aged 65–74, those

75–84, and those 85+. The geographic unit used in this analysis was the county (or county equivalent). Given changes in county definitions during the study period affecting ten counties (e.g., the merging/splitting of counties), a single set of $N_s = 3099$ regions (henceforth referred to simply as counties) from the contiguous lower 48 states (including the District of Columbia) was used for the entire study period. Annual counts of stroke-related deaths per county per age group were obtained from the National Vital Statistics System (NVSS). Due to inconsistencies in the manner in which death records were recorded prior to 1973, the analysis was restricted to data from 1973–2013 ($N_t = 41$ years) to ensure valid comparisons across time. Deaths from stroke were defined as those for which the underlying cause of death was cerebrovascular disease according to the 8th, 9th and 10th revisions of the International Classification of Diseases (ICD; ICD–8: 430–438; ICD–9: 430–438; ICD–10: I60–69). Based on the comparability ratios reported by Klebba and Scott (1980) and Anderson et al. (2001), which indicate a high degree of similarity between the three revisions of the ICD, we assumed that this definition was consistent over the 41-year study period. Annual population counts were based on the bridged-race intercensal estimates provided by NCHS (2013), and we include the percentages of the population that are black and are men as covariates in our analysis based on evidence of disparities in stroke mortality across both race and gender [e.g., Schieb et al. (2013)].

3. Methods. The work of Besag, York and Mollié (1991) has sparked a wealth of research in the disease mapping context for both spatial [e.g., Besag and Higdon (1999), Besag et al. (1995)] and spatiotemporal applications [e.g., Knorr-Held (2000), Knorr-Held and Besag (1998), Waller et al. (1997)]. While these early examples were based on the standard univariate CAR model, Gelfand and Vounatsou (2003) developed methods for general MCAR models, inspiring novel approaches for both multiple and spatiotemporal disease mapping [e.g., Botella-Rocamora, Martinez-Beneito and Banerjee (2015), Martinez-Beneito (2013), Quick, Carlin and Banerjee (2015), Quick, Waller and Casper (2017a)].

3.1. *Statistical model.* Letting Y_{ikt} denote the number of deaths in county i during year t for age group k from a population of size n_{ikt} , we model

$$(1) \quad Y_{ikt} \sim \text{Pois}(n_{ikt} \exp[\mathbf{x}_{ikt}^T \boldsymbol{\beta}_{kt} + Z_{ikt} + \phi_{ikt}])$$

for $i = 1, \dots, N_s$, $k = 1, \dots, N_g$, and $t = 1, \dots, N_t$, where \mathbf{x}_{ikt} denotes a p -vector of covariates with corresponding regression coefficients, $\boldsymbol{\beta}_{kt}$, Z_{ikt} is a spatiotemporal random effect that also accounts for between age-group correlation, and $\phi_{ikt} \stackrel{\text{ind}}{\sim} N(0, \tau_k^2)$. Ignoring the multivariate and temporal sources of correlation momentarily, we could incorporate spatial dependence in our model by following Besag, York and Mollié (1991) and letting

$$(2) \quad Z_{ikt} | \mathbf{Z}_{(i)kt}, \sigma_{kt}^2 \sim N\left(\sum_{j=1}^{N_s} w_{ij} Z_{jkt} / \sum_{j=1}^{N_s} w_{ij}, \sigma_{kt}^2 / \sum_{j=1}^{N_s} w_{ij}\right),$$

where $\mathbf{Z}_{(i)kt}$ denotes the vector $\mathbf{Z}_{\cdot kt} = (Z_{1kt}, \dots, Z_{N_s kt})^T$ with the i th element removed and $w_{ij} = 1$ if i and j are neighbors (denoted $i \sim j$) and 0 otherwise. Recommendations for prior distributions for σ_{kt}^2 and τ_k^2 are offered by [Bernardinelli, Clayton and Montomoli \(1995\)](#).

To account for the multivariate spatiotemporal association in the data, we follow the MSTCAR model of [Quick, Waller and Casper \(2017a\)](#)—itself a special case of the improper MCAR of [Gelfand and Vounatsou \(2003\)](#)—and let $\mathbf{Z} = (\mathbf{Z}_{1..}^T, \dots, \mathbf{Z}_{N_s..}^T)^T \sim \text{MCAR}(1, \boldsymbol{\Sigma}_Z)$, where $\mathbf{Z}_{i..} = (\mathbf{Z}_{i..1}, \dots, \mathbf{Z}_{i..N_t})^T$, $\mathbf{Z}_{i..t} = (Z_{i1t}, \dots, Z_{iN_g t})^T$, $m_i = \sum_{j=1}^{N_s} w_{ij}$, and $\boldsymbol{\Sigma}_Z$ denotes the $N_t N_g \times N_t N_g$ covariance structure for our N_t years and N_g age groups. Full details for constructing $\boldsymbol{\Sigma}_Z$ are provided in Appendix A.1 [[Quick, Waller and Casper \(2017b\)](#)]. Briefly, we begin by defining $\mathbf{v}_{\cdot t} \stackrel{\text{i.i.d.}}{\sim} N(\mathbf{0}, \mathbf{G}_t)$ to be a collection of independent N_g -dimensional random variables with covariance \mathbf{G}_t for $t = 1, \dots, (N_s - 1)$ and $t = 1, \dots, N_t$. Because the MSTCAR model of [Quick, Waller and Casper \(2017a\)](#) is based on an *improper* MCAR model, its spatial domain has dimension $N_s - 1$, hence the range of the subscript t . To account for temporal correlation, we assume $\mathbf{R}_k \equiv \mathbf{R}(\cdot, \cdot | \rho_k)$ denotes the temporal correlation matrix based on an autoregressive order 1 [AR(1)] model for age group k and let $\tilde{\mathbf{R}}_k$ be the Cholesky decomposition of \mathbf{R}_k such that $\tilde{\mathbf{R}}_k \tilde{\mathbf{R}}_k^T = \mathbf{R}_k$. We then define $\boldsymbol{\eta}_{ik\cdot} = \tilde{\mathbf{R}}_k \mathbf{v}_{ik\cdot}$, where $\mathbf{v}_{ik\cdot} = (v_{ik1}, \dots, v_{ikN_t})^T$. Note that if $\mathbf{G}_t \equiv \text{diag}(\sigma_1^2, \dots, \sigma_{N_g}^2)$, then each $\boldsymbol{\eta}_{ik\cdot} \sim N(\mathbf{0}, \sigma_k^2 \mathbf{R}_k)$ is simply an independent AR(1) process with variance σ_k^2 ; allowing these \mathbf{G}_t to be unstructured permits the flexibility for between-group covariances that can evolve over time. We then define $\boldsymbol{\Sigma}_Z$ to be the $N_g N_t \times N_g N_t$ covariance matrix of $\boldsymbol{\eta}_{i..}$ such that $\boldsymbol{\eta}_{i..} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_Z)$. Finally, \mathbf{Z} is constructed from the $\boldsymbol{\eta}_{i..}$ using the eigenvalues and eigenvectors of the adjacency matrix, $\mathbf{W} = \{w_{ij}\}$ [see [Rue and Held \(2005\)](#)]. Following the notation of [Quick, Waller and Casper \(2017a\)](#), we let $\mathcal{G} = \{\mathbf{G}_1, \dots, \mathbf{G}_{N_t}\}$ and $\mathcal{R} = \{\mathbf{R}_1, \dots, \mathbf{R}_{N_g}\}$, and let $\mathbf{Z} \sim \text{MSTCAR}(\mathcal{G}, \mathcal{R})$ denote the model

$$(3) \quad \pi(\mathbf{Z} | \mathcal{G}, \mathcal{R}) \propto |\boldsymbol{\Sigma}_Z|^{-(N_s-1)/2} \exp\left[-\frac{1}{2} \mathbf{Z}^T \{(\mathbf{D} - \mathbf{W}) \otimes \boldsymbol{\Sigma}_Z^{-1}\} \mathbf{Z}\right].$$

3.2. Hierarchical model and computational details. While the Poisson model in (1) is a straightforward extension of the space-only model of [Besag, York and Mollié \(1991\)](#), such models can also pose computational challenges, particularly for large dimensions. For instance, the full conditional of $\mathbf{Z}_{i..}$, given by

$$\begin{aligned} &\pi(\mathbf{Z}_{i..} | \mathbf{Y}, \mathbf{Z}_{(i)..}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\Sigma}_Z) \\ &\propto \prod_{k,t} \text{Pois}(Y_{ikt} | n_{ikt} \exp[\mathbf{x}_{ikt}^T \boldsymbol{\beta}_{kt} + Z_{ikt} + \phi_{ikt}]) \times \pi(\mathbf{Z}_{i..} | \mathbf{Z}_{(i)..}, \boldsymbol{\Sigma}_Z), \end{aligned}$$

is *not* a known distribution; that is, if we use a Markov chain Monte Carlo (MCMC) algorithm to estimate the posterior distribution of our model parameters, this model may require the use of large multivariate Metropolis updates

within our Gibbs sampler. Besag et al. (1995) and Knorr-Held and Rue (2002) suggest a reparameterization of (1) which involves integrating ϕ_{ikt} out of the model, yielding a Gaussian full conditional for $\mathbf{Z}_{i\cdot}$ and requiring Metropolis updates for $\theta_{ikt} = \mathbf{x}_{ikt}^T \boldsymbol{\beta}_{kt} + Z_{ikt} + \phi_{ikt}$. Fortunately, the full conditional distribution of θ_{ikt} —as shown in Web Appendix A.3 [Quick, Waller and Casper (2017b)]—is free of the elements of $\boldsymbol{\theta} \setminus \theta_{ikt}$, thus these Metropolis updates can be conducted independently and in parallel.

We complete our hierarchical model by specifying the following prior distributions for the remaining model parameters: a vague $N(\mathbf{0}, 100 \times \mathbf{I}_p)$ prior for each $\boldsymbol{\beta}_{kt}$, a weakly informative inverse gamma prior for each τ_k^2 , a beta prior for each ρ_k , and an inverse Wishart prior for each \mathbf{G}_t with hyperparameter \mathbf{G} , itself modeled using a vague Wishart prior with diagonal matrix \mathbf{G}_0 . While this hierarchical structure on the covariance matrices is likely unnecessary given the number of spatial regions in the data [see the discussion of prior sensitivity in spatial models by Bernardinelli, Clayton and Montomoli (1995)], this comes at little to no computational cost (see Web Appendix A.2 [Quick, Waller and Casper (2017b)]) and offers a convenient means for specifying proper priors. Putting these pieces together, our joint posterior distribution is as follows:

$$\begin{aligned} \pi(\boldsymbol{\beta}, \mathbf{Z}, \mathbf{G}, \mathcal{G}, \mathcal{R}, \{\tau_k^2\}, \boldsymbol{\theta} | \mathbf{Y}) &\propto \prod_{i,k,t} \text{Pois}(Y_{ikt} | n_{ikt} \exp[\theta_{ikt}]) \times N(\boldsymbol{\theta} | \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}, \boldsymbol{\Sigma}_\theta) \\ &\times \text{MSTCAR}(\mathbf{Z} | \mathcal{G}, \mathcal{R}) \times N(\boldsymbol{\beta} | \mathbf{0}, \boldsymbol{\Sigma}_\beta) \\ (4) \quad &\times \prod_{t=1}^{N_t} \text{InvWish}(\mathbf{G}_t | \mathbf{G}, \nu) \times \text{Wish}(\mathbf{G} | \mathbf{G}_0, \nu_0) \\ &\times \prod_{k=1}^{N_g} [\text{Beta}(\rho_k | a_\rho, b_\rho) \times \text{IG}(\tau_k^2 | a_\tau, b_\tau)], \end{aligned}$$

where $\boldsymbol{\Sigma}_\theta$ is a diagonal matrix of size $N_s N_g N_t$ with elements τ_k^2 , $\boldsymbol{\Sigma}_\beta = 100 \mathbf{I}_{p N_g N_t}$, and \mathbf{X} is the $(N_s N_g N_t \times p)$ matrix of covariates. Additional structure (e.g., temporal correlation) may be considered for $\boldsymbol{\beta}$ as needed. Model comparison will be measured via the deviance information criterion [DIC; Spiegelhalter et al. (2002)]. As a diagnostic, spatial autocorrelation in the posterior estimates of $\lambda_{\cdot kt}$ is summarized by Moran's I index [Moran (1950)]:

$$I(\lambda_{\cdot kt}) = \frac{N_s}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (\lambda_{ikt} - \bar{\lambda}_{\cdot kt})(\lambda_{jkt} - \bar{\lambda}_{\cdot kt})}{\sum_i (\lambda_{ikt} - \bar{\lambda}_{\cdot kt})^2},$$

where $\bar{\lambda}_{\cdot kt} = \sum_i \lambda_{ikt} / N_s$.

While full details for implementing this model in an MCMC framework are provided in Web Appendix A [Quick, Waller and Casper (2017b)], we would be remiss to not discuss the computational burden associated with fitting a nonseparable model as opposed to a separable model, that is, letting $\rho_k = \rho$ for all k and

$\mathbf{G}_t = \mathbf{G}$ for all t corresponds to a model with $\boldsymbol{\Sigma}_Z = \mathbf{R}(\cdot, \cdot | \rho) \otimes \mathbf{G}$. First note that by using an AR(1) model for time, we can compute the $\tilde{\mathbf{R}}_{t,t}^*$ elements of $\boldsymbol{\Sigma}_Z$ in closed form, reducing the burden of computing $\boldsymbol{\Sigma}_Z^{-1}$ from an $N_t N_g \times N_t N_g$ matrix inversion to a series of $N_g \times N_g$ matrix inversions, a property shared with the separable formulation. Furthermore, while the nonseparable MSTCAR model contains more parameters than its separable counterpart, the additional computational burden associated with its implementation in an MCMC framework is negligible; more specifically, the steps required to construct the full conditional distribution of \mathbf{G} in a separable model can each be applied to a specific \mathbf{G}_t in (4).

4. Analysis of the stroke mortality data. We analyzed the stroke mortality data described in Section 2 using the hierarchical model in (4) with the percent nonwhite and percent male as covariates. Due to small population sizes in many counties, the covariate information was smoothed prior to the analysis; see Web Appendix A.5 [Quick, Waller and Casper (2017b)] for more details. The analysis was conducted using three chains of 10,000 iterations, the first 5000 of which were discarded as burn-in. When running the MCMC algorithm, we thinned our posterior samples for θ_{ikt} by removing 9 out of 10 samples—while this is not theoretically necessary, it reduced the burden of storing excess samples for our nearly 400,000 rate parameters. Estimates provided are based on posterior medians, and 95% credible intervals (95% CI) were obtained by finding the 2.5- and 97.5-percentiles from the thinned post-burn-in samples. Additional figures, including animations displaying temporal evolutions in the geographic trends, can be found in Web Appendix B [Quick, Waller and Casper (2017b)].

To determine if the flexibility of the MSTCAR model is necessary for these data, we compared the DIC from the MSTCAR to a collection of N_g independent spatiotemporal processes (i.e., a model which restricts each \mathbf{G}_t to be a diagonal matrix) and a separable multivariate space–time CAR model (i.e., a model which restricts $\mathbf{G}_t = \mathbf{G}$ and $\rho_k = \rho$). As shown in Table 1, the MSTCAR model outperforms both the independent S-T model and the separable model. This is likely due

TABLE 1

DIC comparison between the MSTCAR model, a separable multivariate space–time CAR model and a collection of N_g independent spatiotemporal processes. Smaller values indicate a better compromise between model fit and model complexity. For convenience, differences between the MSTCAR and the competing models are shown in parentheses

Age Group	MSTCAR	Separable	Independent S-T
65–74	16,221,366	16,221,854 (+488)	16,222,781 (+1415)
75–84	26,917,887	26,918,207 (+320)	26,921,159 (+3272)
85+	22,766,145	22,766,806 (+661)	22,768,093 (+1948)
Overall	65,905,398	65,906,867 (+1469)	65,912,033 (+6635)

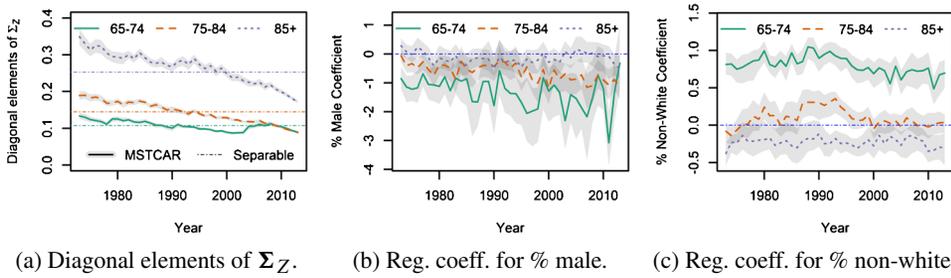


FIG. 1. Parameter estimates from the analysis of the stroke mortality data. Panel (a) compares the diagonal elements of Σ_Z from the MSTCAR model to those from the separable model. Panels (b) and (c) display the temporal evolution of the regression coefficients for the percent of the population that is male and nonwhite, respectively.

to its ability to both account for dependence between the age groups and allow for temporal evolution in the covariance matrix, Σ_Z . As demonstrated in Figure 1(a), the diagonal elements of Σ_Z vary substantially over time. One result of this is that the separable model will tend to *oversmooth* estimates early in the study period (i.e., where the separable model underestimates the elements of Σ_Z) and *undersmooth* estimates later in the study period (i.e., when the separable model overestimates the elements of Σ_Z).

We now dig deeper into the results of the analysis of the stroke mortality data using the MSTCAR model. Figure 1(b) and (c) display the regression coefficients for percent male and percent nonwhite over time for each age group. While these effects do not change substantially between 1973 and 2013, their effects do diminish as age increases. For instance, while there is a positive association between the percent of the population that is nonwhite and increased stroke mortality rates for the 65–74 age group throughout the time period, this trend is less apparent for the 75–84 age group and reversed for those aged 85 and older. This result coincides with the existing literature which suggests that racial disparities in stroke mortality rates are more prominent among younger populations [e.g., Howard (2013)].

Turning our attention to the geographic patterns in stroke death rates, we find substantial differences between age groups (Figures 2–4). For the youngest subpopulation (ages 65–74), the clear geographic patterns shown in Figure 2 highlight the so-called “stroke belt” in the rates from 1973 and the “shift” identified by Casper et al. (1995) toward higher than average rates in parts of Texas and Oklahoma. Similar patterns are also evident in the stroke death rates for those aged 75–84, though the spatial clustering here is less concentrated in the Deep South, with elevated rates spreading into parts of Illinois, Indiana and Michigan. In contrast, the rates for those aged 85+ (Figure 4) exhibit far less spatial clustering with slower rates of decline nationwide. More specifically, this work illustrates that the changing geographic trends identified by Casper et al. (1995) do *not* apply uniformly to all US residents 65 and older, but rather appear most prominently in the

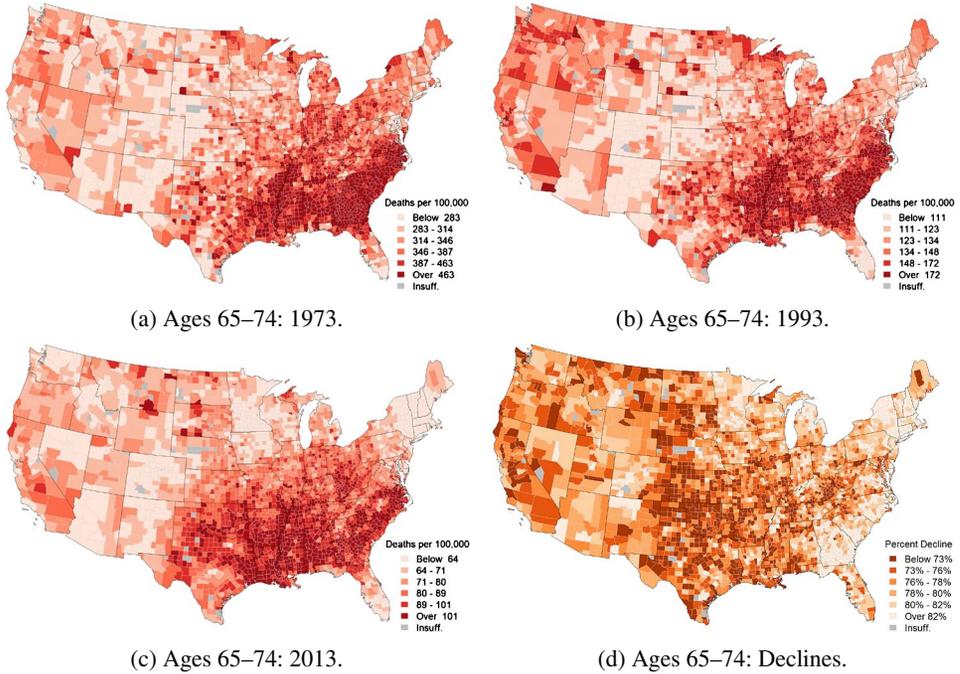


FIG. 2. Maps of the stroke mortality rates and declines for those aged 65–74. Note that estimates for counties with fewer than 100 people in an age bracket in 1973 are suppressed.

youngest age group of our study, and reinforces the need for spatiotemporal analyses of such datasets. See Figures B.3–B.5 of Web Appendix B [Quick, Waller and Casper (2017b)] for map animations for all three age groups and for all 41 years.

Reaffirming our earlier suspicions, Figure 5(a) displays the temporal evolution of the Moran’s I for λ_{ikt} for each age group. Here, we see that the Moran’s I corresponding to the death rates for those aged 65–74 is significantly higher than the remaining age groups with progressively lower degrees of spatial autocorrelation present for the older ages. One caveat to these results is that the 65–74 age group historically has the lowest counts (see Figure B.1), and thus our posterior estimates for λ_{ikt} may have more reliance on the spatial structure in the model. Nonetheless, while overall death counts have declined nearly 50% for both the 65–74 and 75–84 age groups and remained relatively constant for those aged 85 and older, we observe a 25% increase in the Moran’s I from 1973 to 2013 for those 75–84.

Figure 5(b) displays the temporal evolution in the ratios between the average stroke mortality rate in the highest sextile (see Figures 2–4) and the average rate in the lowest sextile by age group. While all three age groups begin the study with ratios on the order of 2:1, the ratios for those aged 75–84 and those 85+ experience significant declines during the 41-year study period. In contrast, the ratio for the 65–74 age group declines at a steady rate from 1973 to the late 1990s before a sharp increase back to a 2.2:1 ratio. Because the highest rates among those aged

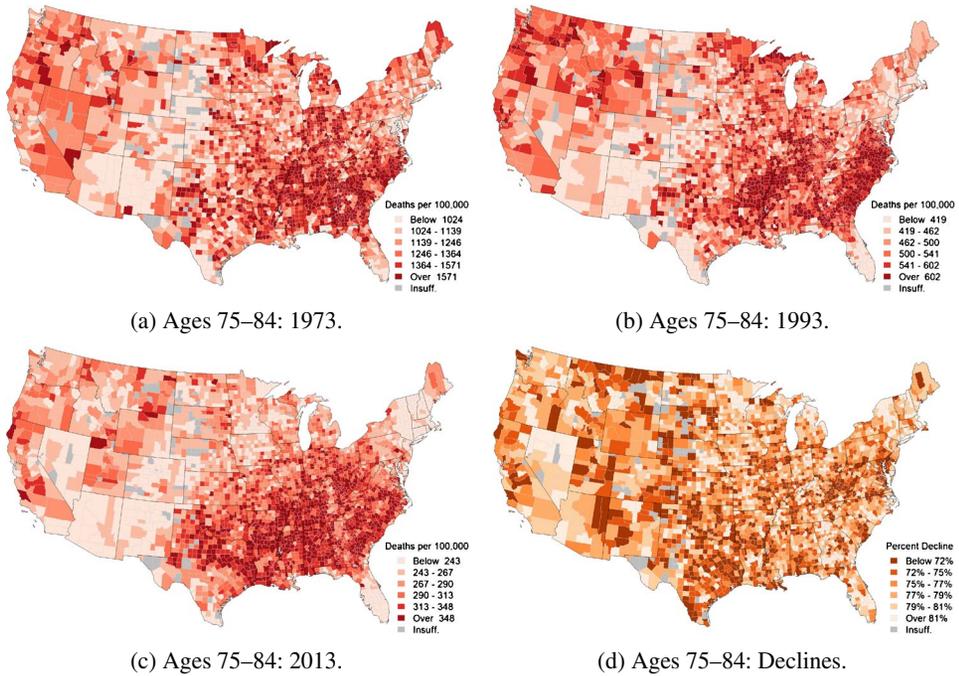


FIG. 3. Maps of the stroke mortality rates and declines for those aged 75-84. Note that estimates for counties with fewer than 100 people in an age bracket in 1973 are suppressed.

65-74 occur predominantly in the Deep South, this suggests that the geographic disparity between this region and the rest of the country has yet to improve.

5. Discussion. By applying the MSTCAR model to these data, we have greatly enhanced our knowledge of spatiotemporal trends in stroke mortality beyond that which existed in the literature. For instance, this work has revealed substantial geographic disparities in the reduction of stroke mortality rates by age group, extending the work of Gillum, Kwagyan and Obiesesan (2011) and Schieb et al. (2013). While Casper et al. (1995) first noted the western shift in the stroke belt, analyses of geographic trends in rates of stroke mortality among *subgroups* of the population have not been conducted before, due in part to a lack of methods equipped for multivariate spatiotemporal modeling. Through this work, we have identified not only that the geographic shift in stroke mortality among those aged 65 and older is largely attributable to those aged 65-74, but also that rates of stroke mortality for those aged 65-74 exhibit more spatial clustering with larger geographic disparities than rates for those 75 and older. In our future work, we aim to further explore these patterns by including younger populations (e.g., adults aged 35 and older) and stratifying our analysis by both race and gender. As this will lead to much smaller counts (due to the reduction in population sizes and the lower event rates among younger populations), we believe the flexibility of the

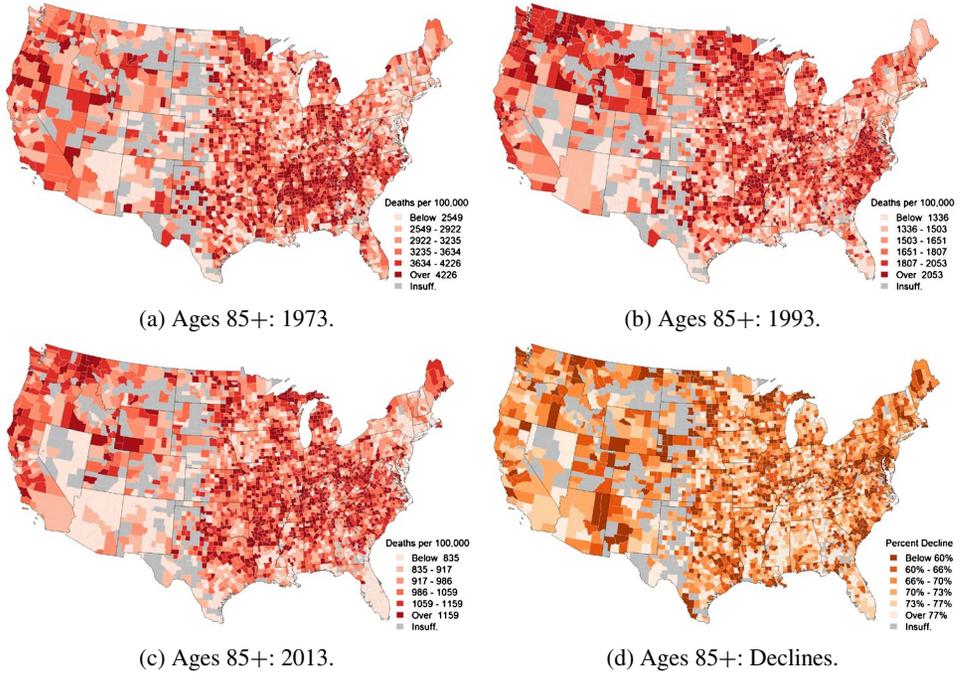


FIG. 4. Maps of the stroke mortality rates and declines for those aged 85+. Note that estimates for counties with fewer than 100 people in an age bracket in 1973 are suppressed.

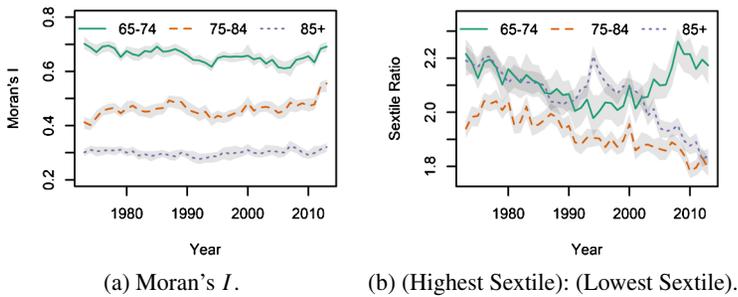


FIG. 5. Panel (a) displays the posterior estimates of the Moran's I for each age group over time. Panel (b) displays the ratio of the average mortality rate for the highest sextile of rates to the average mortality rate for the lowest sextile for each age group.

MSTCAR model makes it a perfect candidate for this analysis and will provide incredible insight into the race/gender disparities in public health.

Another area for future research is understanding the factors that contribute to differential geographic patterns by age group in both the baseline 1973 stroke mortality rates as well as the patterns of declining stroke mortality rates. While it is well known that the risk for stroke increases with age, the spatiotemporal patterns of stroke mortality by age group have not been documented previously. Hypothe-

ses for understanding the observed differential spatiotemporal patterns in stroke mortality by age group include, but are not limited to, the following categories: (a) spatiotemporal differences in the relative contributions of decreasing case fatality rates and incidence rates by age group [e.g., El-Saed et al. (2006), Ergin et al. (2004)]; (b) differential influence of living conditions (e.g., socioeconomic resources, access to quality health care, access to healthy food and recreational environments, etc.), or changes in those living conditions, on stroke mortality by age group [e.g., Lisabeth et al. (2006), Tassone, Waller and Casper (2009)]; or (c) differences in the accuracy of death certificate reporting by age group due to more co-morbidities and competing conditions of death in the older ages.

While there has been and will continue to be a wealth of research related to stroke mortality, the degree to which high quality data are available is an issue which may impede this work. For instance, agencies such as the National Center for Health Statistics (NCHS) within the Centers for Disease Control and Prevention (CDC) are subject to data confidentiality restrictions which require the *suppression* of small counts in public-use data sets [CDC (2003)]. To put such restrictions in context, over 80% of the counts for the 65–74 age group and nearly 70% of the more than 380,000 data points in this analysis are below the recommended threshold of 10 events. Rather than release suppressed data, however, a potential alternative is for agencies to release multiply imputed synthetic data [e.g., Raghunathan, Reiter and Rubin (2003)]. Based on the work here, the MSTCAR model may show promise as a means of synthesizing data for such uses, thereby facilitating future work in this and related areas of research.

Acknowledgments. This work originated while the first author was employed at the Centers for Disease Control and Prevention (CDC). The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the CDC. The authors would like to thank Dr. Michael Kramer and Dr. Adam Vaughan for their comments related to this work.

SUPPLEMENTARY MATERIAL

Supplement to “Multivariate spatiotemporal modeling of age-specific stroke mortality” (DOI: [10.1214/17-AOAS1068SUPP](https://doi.org/10.1214/17-AOAS1068SUPP); .pdf). Appendix A contains the details of our Markov chain Monte Carlo (MCMC) algorithm and a description of the preprocessing smoothing approach used on our two covariates. Appendix B contains a supplemental discussion (and additional figures) related to the analysis of the stroke mortality data.

REFERENCES

- ANDERSON, R. N., MINIÑO, A. M., HOYERT, D. L. and ROSENBERG, H. M. (2001). Comparability of cause of death between ICD-9 and ICD-10: Preliminary estimates. *Natl. Vital Stat. Rep.* **49** 1–32.

- BERNARDINELLI, L., CLAYTON, D. and MONTOMOLI, C. (1995). Bayesian estimates of disease maps: How important are priors? *Stat. Med.* **14** 2411–2431.
- BESAG, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Statist. Soc. Ser. B* **36** 192–236. [MR0373208](#)
- BESAG, J. and HIGDON, D. (1999). Bayesian analysis of agricultural field experiments. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **61** 691–746. [MR1722238](#)
- BESAG, J., YORK, J. and MOLLIE, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Ann. Inst. Statist. Math.* **43** 1–59. [MR1105822](#)
- BESAG, J., GREEN, P., HIGDON, D. and MENGENSEN, K. (1995). Bayesian computation and stochastic systems. *Statist. Sci.* **10** 3–66. [MR1349818](#)
- BORHANI, N. O. (1965). Changes and geographic distribution of mortality from cerebrovascular disease. *Am. J. Publ. Health* **55** 673–681.
- BOTELLA-ROCAMORA, P., MARTINEZ-BENEITO, M. A. and BANERJEE, S. (2015). A unifying modeling framework for highly multivariate disease mapping. *Stat. Med.* **34** 1548–1559. [MR3334675](#)
- CASPER, M., ANDA, R. F., KNOWLES, M., POLLARD, R. A. and WING, S. (1995). The shifting stroke belt: Changes in the geographic pattern of stroke mortality in the United States. *Stroke* **26** 755–760.
- CDC (2003). CDC/ATSDR Policy on Releasing and Sharing Data. Manual; Guide CDC-02. Available at <http://www.cdc.gov/maso/Policy/ReleasingData.pdf>. Accessed June 30, 2015.
- EL-SAEED, A., KULLER, L. H., NEWMAN, A. B., LOPEZ, O., COSTANTINO, J., MCTIGUE, K., CUSHMAN, M. and KRONMAL, R. (2006). Geographic variations in stroke incidence and mortality among older populations in four US communities. *Stroke* **37** 1975–1979.
- ERGIN, A., MUNTNER, P., SHERWIN, R. and HE, J. (2004). Secular trends in cardiovascular disease mortality, incidence, and case fatality rates in adults in the United States. *Am. J. Med.* **117** 219–227.
- GELFAND, A. E. and VOUNATSOU, P. (2003). Proper multivariate conditional autoregressive models for spatial data analysis. *Biostat.* **4** 11–25.
- GILLUM, R. F., KWAGYAN, J. and OBISESAN, T. O. (2011). Ethnic and geographic variation in stroke mortality trends. *Stroke* **42** 3294–3296.
- HOWARD, V. J. (2013). Reasons underlying racial differences in stroke incidence and mortality. *Stroke* **44** S126–S128.
- HOWARD, G., HOWARD, V. J., KATHOLI, C., OLI, M. K., HUSTON, S. and ASPLUND, K. (2001). Decline in US stroke mortality: An analysis of temporal patterns by sex, race, and geographic region. *Stroke* **32** 2213–2218.
- KLEBBA, A. J. and SCOTT, J. H. (1980). Estimates of selected comparability ratios based on dual coding of 1976 death certificates by the eighth and ninth revisions of the International Classification of Diseases. *National Vital Statistics Reports* **28**.
- KNORR-HELD, L. (2000). Bayesian modelling of inseparable space–time variation in disease risk. *Stat. Med.* **19** 2555–2567.
- KNORR-HELD, L. and BESAG, J. (1998). Modelling risk from a disease in time and space. *Stat. Med.* **17** 2045–2060.
- KNORR-HELD, L. and RUE, H. (2002). On block updating in Markov random field models for disease mapping. *Scand. J. Stat.* **29** 597–614. [MR1988414](#)
- LISABETH, L. D., DIEZ ROUX, A. V., ESCOBAR, J. D., SMITH, M. A. and MORGENSTERN, L. B. (2006). Neighborhood environment and risk of ischemic stroke. *Am. J. Epidemiol.* **165** 279–287.
- MARTINEZ-BENEITO, M. A. (2013). A general modelling framework for multivariate disease mapping. *Biometrika* **100** 539–553. [MR3094436](#)
- MORAN, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika* **37** 17–23. [MR0035933](#)

- NCHS (2013). Bridged-race population estimates: United States July 1st resident population by state, county, age, sex, bridged-race, and Hispanic origin. Available at http://www.cdc.gov/NCHS/nvss/bridged_race.htm. Accessed June 30, 2015.
- QUICK, H., CARLIN, B. P. and BANERJEE, S. (2015). Heteroscedastic conditional auto-regression models for areally referenced temporal processes for analysing California asthma hospitalization data. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **64** 799–813. [MR3415951](#)
- QUICK, H., WALLER, L. A. and CASPER, M. (2017a). A multivariate space–time model for analyzing county-level heart disease death rates by race and sex. *J. Roy. Statist. Soc. Ser. C.* DOI:[10.1111/rssc.12215](#).
- QUICK, H., WALLER, L. A. and CASPER, M. (2017b). Supplement to “Multivariate spatiotemporal modeling of age-specific stroke mortality.” DOI:[10.1214/17-AOAS1068SUPP](#).
- RAGHUNATHAN, T. E., REITER, J. P. and RUBIN, D. B. (2003). Multiple imputation for statistical disclosure limitation. *J. Off. Stat.* **19** 1–16.
- RUE, H. and HELD, L. (2005). *Gaussian Markov Random Fields: Theory and Applications. Monographs on Statistics and Applied Probability* **104**. Chapman & Hall/CRC, Boca Raton, FL. [MR2130347](#)
- SCHIEB, L. J., MOBLEY, L. R., GEORGE, M. and CASPER, M. (2013). Tracking stroke hospitalization clusters over time and associations with county-level socioeconomic and healthcare characteristics. *Stroke* **44** 146–152.
- SPIEGELHALTER, D. J., BEST, N. G., CARLIN, B. P. and VAN DER LINDE, A. (2002). Bayesian measures of model complexity and fit. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64** 583–639. [MR1979380](#)
- TASSONE, E. C., WALLER, L. A. and CASPER, M. L. (2009). Small-area racial disparity in stroke mortality. *Epidemiology* **20** 234–241.
- WALLER, L. A., CARLIN, B. P., XIA, H. and GELFAND, A. E. (1997). Hierarchical spatio-temporal mapping of disease rates. *J. Amer. Statist. Assoc.* **92** 607–617.
- XU, J., MURPHY, S. L., KOCHANNEK, K. D. and BASTIAN, B. A. (2016). Deaths: Final data for 2013. *Natl. Vital Stat. Rep.* **64** 1–119.

H. QUICK
DEPARTMENT OF EPIDEMIOLOGY
AND BIOSTATISTICS
DREXEL UNIVERSITY
PHILADELPHIA, PENNSYLVANIA 19104
USA
E-MAIL: hsq23@drexel.edu

L. A. WALLER
DEPARTMENT OF BIostatISTICS
AND BIOINFORMATICS
EMORY UNIVERSITY
ATLANTA, GEORGIA 30322
USA
E-MAIL: lwaller@emory.edu

M. CASPER
DIVISION OF HEART DISEASE
AND STROKE PREVENTION
CENTERS FOR DISEASE CONTROL
AND PREVENTION
ATLANTA, GEORGIA 30329
USA
E-MAIL: myc5@cdc.gov