Modelling Risk from a Disease in Time and Space

Leonhard Knorr-Held **Julian Besag**

NRCSE

Technical Report Series

NRCSE-TRS No. 005

Modelling Risk from a Disease in Time and Space

Leonhard Knorr–Held*
Institut für Statistik
Universität München
Ludwigstr. 33, 80539 München
Germany

Julian Besag
Department of Statistics
University of Washington
Seattle, WA 98195
U.S.A.

May 27, 1997

Abstract

This paper combines existing models for longitudinal and spatial data in a hierarchical Bayesian framework, with particular emphasis on the role of time– and space–varying covariate effects. Data analysis is implemented via Markov chain Monte Carlo methods. The methodology is illustrated by a tentative re-analysis of Ohio lung cancer data 1968-88. Two approaches that adjust for unmeasured spatial covariates, particularly tobacco consumption, are described. The first includes random effects in the model to account for unobserved heterogeneity; the second adds a simple urbanization measure as a surrogate for smoking behaviour. The Ohio dataset has been of particular interest because of the suggestion that a nuclear facility in the southwest of the state may have caused increased levels of lung cancer there. However, we contend here that the data are inadequate for a proper investigation of this issue.

*Email: leo@stat.uni-muenchen.de

1 Introduction

Data on disease incidence or mortality over a set of contiguous geographical regions are very common. Typically such data are stratified according to several covariates, with the number of persons and the number of cases or deaths being recorded for each region and covariate combination. There can be several reasons for a spatial analysis of such data. For example, a risk map may be required as an aid to the allocation of resources or as an attempt to discover new factors regarding the disease. Raw or standardized incidence rates may produce very unreliable maps, because some populations at risk are small and/or the disease is rather rare. The goal here could be described as one of smoothing, in which both spatial and non-spatial considerations may arise; the examination of residuals may be particularly important. At the other end of the spectrum, there may be a quite concrete task, perhaps involving the assessment of additional risk from a known putative source, such as a nuclear installation. For some further comments contrasting "focussed" and "unfocussed" analysis in a frequentist setting, see Besag and Newell (1991). For Bayesian methodology and examples, mainly with regard to unfocussed analysis, see Clayton and Kaldor (1987), Besag, York and Mollié (1991), Clayton, Bernardinelli and Montomoli (1993) and, more comprehensively, Clayton and Bernardinelli (1993).

Longitudinal information on incidence rates introduces a further dimension that is often crucially important because of changes in socio-demographic structure or other risk factors during the time period of the study. In the present paper, we use for illustration a dataset on mortality from lung cancer for 21 successive years in the 88 counties of Ohio. We describe the data in Section 2 of the paper. They have been analyzed previously by Waller, Carlin, Xia and Gelfand (1997) and have been of particular interest because of the suggestion that a nuclear installation in the southwest of the state may have been responsible for an increased level of lung cancer there; for a purely spatial analysis, see Devine (1992) and Devine, Louis and Halloran (1994). Although we comment further on this issue during the course of the paper, we contend that Bayesian smoothing is not the most appropriate tool for focussed analysis and that, in any case, the Ohio dataset does not provide enough information for any proper conclusion to be drawn. Thus, our main concern with these data is as an illustration of Bayesian mapping in time and space and, even in this limited pursuit, we do not yet

produce a complete solution.

Among the many formulations developed for non-spatial analysis of longitudinal data, hierarchical dynamic models play a central role in our approach. Such models accommodate an appropriate degree of temporal flexibility and can be easily combined with spatial formulations already in use for the analysis of cross-sectional data. Perhaps the most compelling feature of dynamic models is that the parameters associated with time trends and specific covariates are free to vary over time; see, for example, Fahrmeir and Tutz (1994, Ch.8). Thus, they provide a generalization of the state space models that have been popular in analyzing approximately Gaussian data (e.g. Harvey, 1989; West and Harrison, 1989). Dynamic models in a Bayesian framework aim to produce smoothness in time trends, so that estimates for any particular time can "borrow strength" from data at adjacent times. The corresponding prior distributions neither impose stationarity nor assume a specific parametric form; in fact dynamic models are related to the semi– and non–parametric smoothing methods of Hastie and Tibshirani (1990), as reviewed by Fahrmeir and Knorr-Held (1997a). Recent applications of dynamic models for longitudinal and survival data are given in Berzuini and Clayton (1994), Besag, Green, Higdon and Mengersen (1995), Berzuini and Larizza (1996) and Fahrmeir and Knorr–Held (1997b).

In Section 3, we describe our approach to time-space modelling of disease risk data, in the particular context of the Ohio dataset, though it applies more widely. We also include some details of the Markov chain Monte Carlo (MCMC) computations used in implementing the corresponding statistical analysis. In Section 4, we discuss the numerical results for the Ohio data, both in applying the basic model and an extended version that attempts to capture smoking patterns more explicitly via the Kafadar–Tukey index. Section 5 provides some general discussion.

2 Ohio lung cancer data

The state of Ohio is located in the northeast of the United States and is divided into 88 counties. The database records the population size and the number of deaths from lung cancer, stratified by age, gender and race (white or non–white), for each year between 1968

and 1988 and for each county. Age is categorized as under 15 years, 15–24, 25–34, ..., 65–74, and over 74 years. Since very few deaths from lung cancer occur below age 45, we combine the first four categories, which therefore results in five agegroups, labelled 1 through 5.

The spatial variability in the crude annual death rates $\times 1000$ is shown in Figure 1. This map does not allow for differential risk factors The variability over time is illustrated in Figure 2, which focuses on 45 to 54 year–old white males in four different counties and displays time series of the numbers at risk, the numbers of deaths and the corresponding mortality rates. As one might expect, the widest range in the mortality rates occurs in the two counties with the lowest populations at risk, Lawrence and Wyondot.

Figure 3 shows the observed mortality rate from lung cancer over time, broken down by gender and race in the top panel, ignoring deaths below age 45, and by agegroup in the bottom one. However, these plots can be misleading. For example, we must not conclude that there is a higher risk for white than for non–white women, since the difference in the mortality rates might be an artifact of disparities in the age distributions of the two subgroups. Indeed, later we shall see that this appears to be the case.

Of course, a major risk factor for lung cancer is cigarette consumption and any analysis that ignores this is highly questionable. Our basic formulation allows for unobserved covariates in each county but this device is unlikely to be entirely satisfactory here. Although there is no reliable direct information on smoking behaviour across Ohio, it should help to include a measure of urbanization in the model, primarily as a surrogate for cigarette consumption. For example, Kafadar and Tukey (1993) suggest the logarithm of the population size of the largest city in each county and compare this favourably with other, more obvious, measures, such as population density, in a nationwide analysis of deaths from lung cancer. Figure 4 shows a map of the Kafadar–Tukey (K–T) index for Ohio, based on population data from 1970. The four counties containing the major cities (Cleveland, on Lake Erie, in the north; Cincinnati, in the southwest; Columbus, approximately in the middle of the state; and Toledo, on the Michigan/Ohio border, in the northwest) are clearly identified.

3 Time–space modelling

We describe our model formulation in terms of the Ohio dataset but of course it applies more generally with appropriate modification. Thus, let n_{ijkt} denote the number at risk in county i ($i = 1, \ldots, I$) and year t ($t = 1, \ldots, T$), for a specific agegroup j ($j = 1, \ldots, J$) and gender $\times$ race combination k ($k = 1, \ldots, K$); here $I = 88$, $T = 21$, $J = 5$ and $K = 4$. We assume that the number of deaths y_{ijkt} from lung cancer, during year t and in category (i, j, k), has a binomial distribution with parameters n_{ijkt} and π_{ijkt}, and that the likelihood for the entire data is the corresponding product of binomial terms. In some contexts, a Poisson approximation to the binomial might be appropriate but this has no real advantage and here there are some categories for which the death rates are uncomfortably large.

The above binomial formulation might be adopted as a convenient approximation in a frequentist framework but there is also a subjective justification. Thus, consider a particular individual who is in category (i, j, k) during year t. Suppose that she belongs to some finer risk category l; for example, she may be a heavy smoker, working in the steel industry. Now assume that she will die of lung cancer during the year with a probability that is a random draw from a distribution with mean π_{ijklt}; this distribution might or might not reduce to a single atom. Then, under independence, the number of deaths in category (i, j, k, l, t) is binomial with parameters n_{ijklt}, the relevant number at risk, and π_{ijklt}. However, the n_{ijklt} are unknown to us and we therefore suppose they form a multinomial sample in which an (i, j, k, t) individual falls in group l with probability ν_{ijklt}. It follows that y_{ijkt} has a binomial distribution of the required form, where

$$\pi_{ijkt} = \sum_l \nu_{ijklt} \pi_{ijklt}$$

and it is only this probability that we need to model. Note that, in general, π_{ijkt} is not the probability of death for anyone in particular but is merely the mean of a distribution. Of course, this is a weakness but one that cannot be avoided without making additional rather strong assumptions.

We follow a standard path in modelling π_{ijkt} with a logit link to the binomial and use a linear predictor η_{ijkt} that decomposes additively into time– and space–dependent effects. For some motivation, consider a single individual in each category (i, j, k, t) and suppose

that π_{ijkt} *is* the corresponding probability of death from lung cancer. Thus, $n_{ijkt} = 1$ and $y_{ijkt} = 0$ or 1. Then the simplest plausible assumption is that, conditional on a single death, this occurs to the individual in category (i, j, k, t) with a probability that is the product of the corresponding marginals; that is,

$$\Pr(y_{ijkt} = 1 | y_{++++} = 1) = \Pr(y_{i+++} = 1 | y_{++++} = 1) \Pr(y_{+j++} = 1 | y_{++++} = 1)$$

$$\times \Pr(y_{++k+} = 1 | y_{++++} = 1) \Pr(y_{+++t} = 1 | y_{++++} = 1).$$

The probability on the left–hand side is proportional to $\pi_{ijkt}/(1 - \pi_{ijkt})$ and thus we obtain a simple additive model for the logistic transform of π_{ijkt}. However, this simple independence assumption is too naive and we must allow some interactions, though we maintain the separability of space and time. Our eventual model is that the log–odds,

$$\eta_{ijkt} = \ln\{\pi_{ijkt}/(1 - \pi_{ijkt})\},$$

has the decomposition,

$$\eta_{ijkt} = \alpha_t + \beta_{jt} + \gamma_{kt} + \delta z_i + \theta_i + \phi_i, \tag{1}$$

where α_t is the effect of year t, β_{jt} is the agegroup j effect at time t, γ_{kt} is the gender $\times$ race effect for combination k at time t, and z_i is the (centred) K–T index in county i, if included. Finally, θ_i and ϕ_i represent unspecified features of county i that respectively do or do not display spatial structure and can be interpreted as surrogates for unmeasured covariates, as in Besag *et al.* (1991). If information on relevant temporal covariates was not available, we would include surrogates here as well but the model would become very crude.

We have also used a more general formulation, in which interactions between gender, agegroup and time and between race, agegroup and time are allowed in equation (1) but the inclusion of such terms makes little difference here. Note that there is no simple way to include cohort effects, because information on age is not available on an annual basis. In principle, one might remedy this using a missing values formulation, as suggested in Besag *et al.* (1995), but this requires knowledge of the number at risk on an annual basis both for age and calendar time. Here, we have allowed for cohort effects indirectly via the agegroup $\times$ period interactions β_{kt}.

The formulation (1) is completed by assigning prior distributions to the various components of η_{ijkt}. For effects that are a function of time, we do not expect exchangeability and must seek a viable alternative. Thus, for $\alpha_1, \ldots, \alpha_T$, we adopt a random walk with independent Gaussian increments, conditional on a variance σ_α^2 which is then assigned a highly dispersed but proper hyperprior. Note that the overall effect of the η_{ijkt} can be absorbed by the level of the random walk. Other possible priors include those based on second differences, as opposed to first differences, as in Berzuini *et al.* (1993) or Besag *et al.* (1995), and increments that conform to the hierarchical-t structure in Besag and Higdon (1997).

As regards the other temporal effects, we wish to make analogous assumptions but also to avoid identifiability problems. Therefore, we constrain the means $\beta_{.t}$ and $\gamma_{.t}$ to be zero for each t. This device results in singular covariance matrices, which are known apart from the individual variances $\sigma_{\beta_j}^2$ and $\sigma_{\gamma_k}^2$. A related approach is described by Harvey (1989); for full details in the present context, see Knorr-Held (1996). Note that it is not valid here merely to set some parameter values to zero because this would destroy the prior exchangeability of the β_{jt} and the γ_{kt} for each fixed t. Note also that, even though Figure 3 happens not to conflict with linearity, we would not want to impose deterministic time trends on the data nor to make an assumption of stationarity.

For the spatially structured components θ_i, we choose a simple Gaussian intrinsic autoregression; see, for example, Besag *et al.* (1991). Thus, the conditional distribution of θ_i is

$$\theta_i | \theta_{-i}, \sigma_\theta^2 \sim N(\bar{\theta}_i, \sigma_\theta^2/m_i), \qquad i = 1, \ldots, I,$$

where $\bar{\theta}_i$ is the corresponding mean value over the m_i counties that are geographically contiguous to i, and σ_θ^2 is a variance parameter. This autoregression is a spatial analogue of the random walk and similarly is just non-stationary. For an alternative median–related prior, see Besag *et al.* (1991), and for further relevant discussion of intrinsic autoregressions, Besag and Kooperberg (1995). For different choices of neighbourhoods and weights, see Besag (1975) and the rejoinder in Besag *et al.* (1991). Finally, unstructured spatial heterogeneity is accounted for by assuming the ϕ_i to be independent and Gaussian with mean zero and variance σ_ϕ^2. Both σ_θ^2 and σ_ϕ^2 are assigned highly dispersed but proper inverse gamma priors.

We used Markov chain Monte Carlo to sample from the posterior distribution implied

by the above formulation. Specifically, we applied univariate or block Metropolis steps, as described for example in Smith and Roberts (1993) or Besag *et al.* (1995). Pilot runs confirmed the existence of a large negative correlation between the spatial and non–spatial parameters in some counties, which led to extremely slow mixing. We remedied this by updating θ_i and ϕ_i in a single block. The dispersion matrix of the corresponding bivariate Gaussian proposal was chosen roughly to match the second moments, based on the results of the pilot runs. The θ_i were recentred after each cycle, so that the overall effect was absorbed by the α_t; see Besag *et al.* (1995). Finally, for each t, we updated the time–dependent parameters in blocks, again to improve computational efficiency.

4 Results

In this section, we re–analyze the Ohio dataset, using the formulation described in Section 3, both with and without the K–T index z_i in equation (1). First, we consider the temporal parameters: the pictures we show are for the full model but the results are virtually indistinguishable from those when z_i is excluded. Figure 5 displays the posterior medians and 50, 80 and 95% pointwise credible intervals for the overall time trend α_t and for the aggregate of this with the five agegroup effects β_{jt}, all plotted on the same scale. There is some evidence of a change of slope in the overall effect around 1980. The risks in the five agegroups follow the expected ordering and there is a similar ranking as regards the *increase* in risk over the period of the study. Recall that mortality in agegroup 1 is almost negligible.

Figure 6 shows corresponding plots, aggregating the effects γ_{kt} for the four gender $\times$ race combinations with the overall time trend α_t. There are large differences between the effects for men and for women and between white and non-white men. Although the risk for men shows a slight increase over time and is always much greater than for women, the gender effect declines in importance over the period of the study, in accordance with known increases in cigarette consumption by women. Perhaps the most interesting aspect of Figure 6 is the dominance of the median effect for non–white women over that for white women at each time point; indeed the posterior probability that the risk for non–white women is greater at *all* 21 time points is 0.76 and that the average effect dominates is essentially unity.

Although this result again agrees with expectation, recall that it is in apparent conflict with the observed rates in Figure 3. The explanation is that collapsing over age provides an instance of Simpson's paradox. Suppose instead that one were to collapse over time. Then the observed rates for non–white women would be *greater* in every agegroup, especially the first three, despite the fact that the *overall* rate is higher for white women. Note that, for example, in 1968, non-white women comprised 10.9% of the population under 15 but only 5.4% of those over 65; for 1988, the corresponding figures were 14.4% and 7.4%. Thus, there is confounding in the observed rates between race and age, which the model resolves by the inclusion of (time–dependent) race and age effects. In the analysis by Waller *et al.* (1997), there is no allowance for age effects and hence the anomaly in the rates is inherited by the estimated risks. Note that the purely spatial analyses in Clayton and Kaldor (1987), Besag *et al.* (1991) and Devine *et al.* (1994) employ rates that are standardized for age, which would seem to be a minimal requirement.

In Figure 7, we again plot the observed mortality rates for white males aged between 45 and 54 in four of the counties, as in Figure 2, and now compare these with the posterior median and 50, 80 and 95% credible intervals for the estimated risk of death. We also show corresponding 50 and 80% *simultaneous* credible regions, calculated by the method in Besag *et al.* (1995, Section 6.3), which is equivariant to monotone transformations. Despite the highly disparate observed mortality rates, there is no evidence of differences in risk and the same holds true when comparisons with other counties are made. On the other hand, Figure 8 provides a map of the overall fitted annual death rates $\times 1000$ for each county, and this shows large differences across the state. The map does not relate in any useful way to the change in risk for an individual moving from county to county; and this remark would hold also for any corresponding single-period map. However, it is directly comparable to Figure 1, which it resembles quite closely.

An important advantage of a formulation that does not include time $\times$ space interactions, when this is valid, is that the spatial parameters can be interpreted more easily. In particular, we define the *adjusted relative risk* in county i by

$$\mathrm{ARR}_i = \exp(\theta_i + \phi_i),$$

which is automatically calibrated to a common base for all other covariates. Of course, ARR_i

does *not* measure some average risk from lung cancer in county i but it does address the existence of additional risk factors.

Figures 9 and 10 are maps of the posterior medians of the ARRs, omitting and including the K–T index, respectively. We first discuss the basic unadjusted results. The most striking feature in Figure 9 is the concentration of high values in the southwest of the state. Indeed, three of the four highest ARRs occur there: 1.30 in Hamilton, in the extreme southwest; 1.32 in Clermont, immediately to the east; and 1.25 in Butler, immediately to the north. The corresponding 95% credible intervals are (1.26, 1.35), (1.24, 1.41) and (1.19, 1.31). The posterior probability of an above average ARR is essentially unity in any of these three counties. Recall that Hamilton is of particular interest because it contains the nuclear facility at Fernald. However, these results are all suspect because they do not take account of the K–T index and yet there is clearly a visual similarity between Figures 4 and 9; indeed, the corresponding correlation coefficient is 0.42. Note, incidentally that the variability in the posterior distribution of the ARRs reflects population density and so is also broadly in line with Figure 4 but with an inverse shading.

We turn now to the extended analysis, in which the ARRs are also calibrated for urbanicity and hence, crudely, for smoking behaviour. In general, the posterior medians in Figure 10 are shrunk towards unity, as one might expect. In particular, those for Hamilton and Butler are much reduced to 1.07 and 1.15, respectively. However, the posterior median for Clermont increases to 1.41, because of the very low K–T index there. The corresponding 95% credible intervals are (0.98, 1.17), (1.09, 1.21) and (1.31, 1.50); in Hamilton, the probability of an above average ARR decreases to 0.93. The result for Clermont perhaps merits further investigation but the obvious explanation is that its close proximity to Cincinnati induces an above average level of tobacco consumption. Indeed, one might modify the linear predictor (1) to incorporate information about the K–T indices in adjacent counties but the real message here is that surrogate measures must always be treated with some suspicion.

5 Conclusions

The modelling of data in time and space is a challenging task, made all the more difficult in many environmental applications because the spatial units are not arranged on a regular array. The present paper presents some Bayesian methodology, exemplified by an assessment of risk from lung cancer in the state of Ohio, based on annual observations of mortality in each of the 88 counties. Our formulation includes effects for agegroup and gender $\times$ race, through time, and for space. Rigorous analysis is hampered by the absence of direct information on tobacco consumption in each county and we therefore introduce the Kafadar–Tukey urbanization index as a surrogate measure. We also recognize the existence of other unknown county effects and model these by non-specific surrogates. Although this approach is less than ideal, we hope that our formulation provides a useful stepping stone in the development of time–space methodology for the statistical analysis of risk from (non–communicable) diseases and will perhaps encourage the collection of more complete information in the future. We conclude with some comments towards these goals.

There is of course no doubt that a carefully designed case–control study would be much preferred to the observational data with which the present paper is concerned. The results from observational studies may suggest the need for further investigation but can rarely be seen as an end in themselves. Nevertheless, computational advances during the past decade, particularly the adoption of Markov chain Monte Carlo as a standard Bayesian tool, have largely removed the need for models of convenience, so that one can now contemplate more realistic formulations, incorporating features that might hitherto have been ignored. For example, we believe that discrete dynamical models provide a significant advance in describing effects that develop over time.

The construction of complex stochastic models for observational data is likely to remain a mixture of art and science. All models are incorrect and there is often a quite thin line between those that are too naive to be useful and those that are too complicated to be fitted reliably to the available data. We have not followed Waller *et al.* (1997) in applying formal model choice criteria, though we did use standard contingency table methods on the temporal data, as an exploratory tool. We contend that the above authors selected their models from an inappropriate class and that this led to an analysis with which even they

were ill at ease; in particular, none of their potential formulations incorporated agegroup effects or interactions between period and age, race or gender. Instead, presumably all such factors were to be described by the inclusion of time $\times$ space interactions, which breaks the premise that these represent unknown or unmeasured covariates. Also, there is a considerable advantage in interpretation if one can identify a viable model that is separable in time and space. However, we recognize that time $\times$ space interactions may be an important residual feature, in which case our formulation would require appropriate expansion. As regards the spatial surrogates, recall that the ultimate goal is to remove the need for non–specific covariates, so that ideally neither θ nor ϕ should enter into the formulation. However, in practice, lack of available data or insufficient insight will often lead to a compromise solution, as in the present paper.

Acknowledgements

This research was initiated while the first author was visiting the Department of Statistics, University of Washington, Seattle, U.S.A., whose hospitality is gratefully acknowledged. The visit was supported by a grant from the German Academic Exchange Service (DAAD). Partial support for the second author was provided by the National Science Foundation. The authors express thanks to Owen Devine for access to the Ohio lung cancer data, to Norman Breslow for helpful comments on an earlier version of the paper and to Simon Byers for assistance in computational strategies.

References

Berzuini, C. and Clayton, D. (1994). Bayesian analysis of survival on multiple time scales. *Statistics in Medicine*, **13**, 823–838.

Berzuini, C. and Larizza, C. (1996). A unified approach for modelling longitudinal and failure time data, with application in medical monitoring. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **18**, 109–123.

Berzuini, C., Clayton, D. and Bernardinelli, L. (1993). Bayesian inference on the Lexis diagram. *Bulletin of the International Statistical Institute*, **50**, 149–164.

Besag, J. E. (1975). Statistical analysis of non–lattice data. *The Statistician*, **24**, 179–195.

Besag, J. E. and Higdon, D. M. (1997). Bayesian inference for agricultural field experiments. Submitted.

Besag, J. E. and Newell, J. N. (1991). The detection of clusters in rare diseases. *Journal of the Royal Statistical Society A*, **154**, 143–155.

Besag, J. E., Green, P. J., Higdon, D. M. and Mengersen, K. L. (1995). Bayesian computation and stochastic systems (with discussion). *Statistical Science*, **10**, 3–66.

Besag, J. E. and Kooperberg, C. L. (1995). On conditional and intrinsic autoregressions. *Biometrika*, **82**, 733–746.

Besag, J. E., York, J. C. and Mollié, A. (1991). Bayesian image restoration with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics*, **43**, 1–59.

Clayton, D. G. and Bernardinelli, L. (1993). Bayesian methods for mapping disease risks. In: J. Cuzick and P. Elliot (eds), *Small Area Studies in Geographical and Environmental Epidemiology*, Oxford University Press, pp. 205–220.

Clayton, D. G., Bernardinelli, L. and Montomoli, C. (1993). Spatial correlation in ecological analysis. *International Journal of Epidemiology*, **22**, 1193–1202.

Clayton, D. G. and Kaldor, J. (1987). Empirical Bayes estimates of age–standardized relative risks for use in disease mapping. *Biometrics*, **43**, 671–681.

Devine, O. J. (1992). *Empirical Bayes and Constrained Empirical Bayes Methods for Estimating Incidence Rates in Spatially Aligned Areas.* Unpublished Ph.D thesis, Division of Biostatistics, Emory University.

Devine, O. J., Louis, T. A. and Halloran, M. E. (1994). Empirical Bayes estimators for spatially correlated incidence rates. *Environmetrics*, **5**, 381–398.

Fahrmeir, L. and Knorr-Held, L. (1997a). Dynamic and semiparametric models. In: M. Schimek (ed.), *Smoothing and Regression: Approaches, Computation and Application*, New York: John Wiley & Sons. To appear.

Fahrmeir, L. and Knorr-Held, L. (1997b). Dynamic discrete time duration models: estimation via Markov chain Monte Carlo. *Sociological Methodology*. To appear.

Fahrmeir, L. and Tutz, G. (1994a). *Multivariate Statistical Modelling Based on Generalized Linear Models*. New York: Springer–Verlag.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: University Press.

Hastie, T. and Tibshirani, R. (1990). *Generalized Additive Models*. London: Chapman & Hall.

Kafadar, K. and Tukey, J. W. (1993). U.S. cancer rates: a simple adjustment for urbanization. *International Statistical Review*, **61**, 257–281.

Knorr-Held, L. (1996). *Hierarchical Modelling of Discrete Longitudinal Data; Applications of Markov Chain Monte Carlo*. Unpublished Ph.D. thesis, Institut für Statistik, Universität München.

Smith, A. F. M. and Roberts, G. O. (1993). Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods (with discussion). *Journal of the Royal Statistical Society B*, **55**, 3–23.

Waller, L. A., Carlin, B. P., Xia, H. and Gelfand, A. E. (1997). Hierarchical spatio–temporal mapping of disease rates. *Journal of the American Statistical Association*. To appear.

West, M. and Harrison, P. J. (1989). *Bayesian Forecasting and Dynamic Models*. New York: Springer.

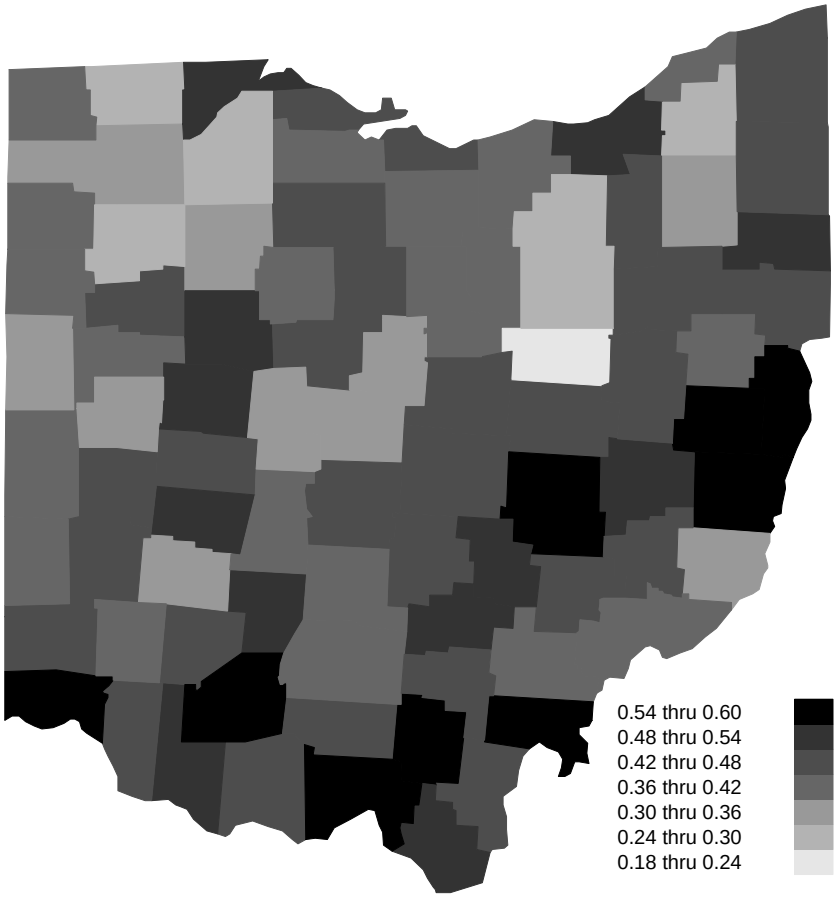

Figure 1: Crude annual death rate $\times$ 1000 for each county in Ohio.

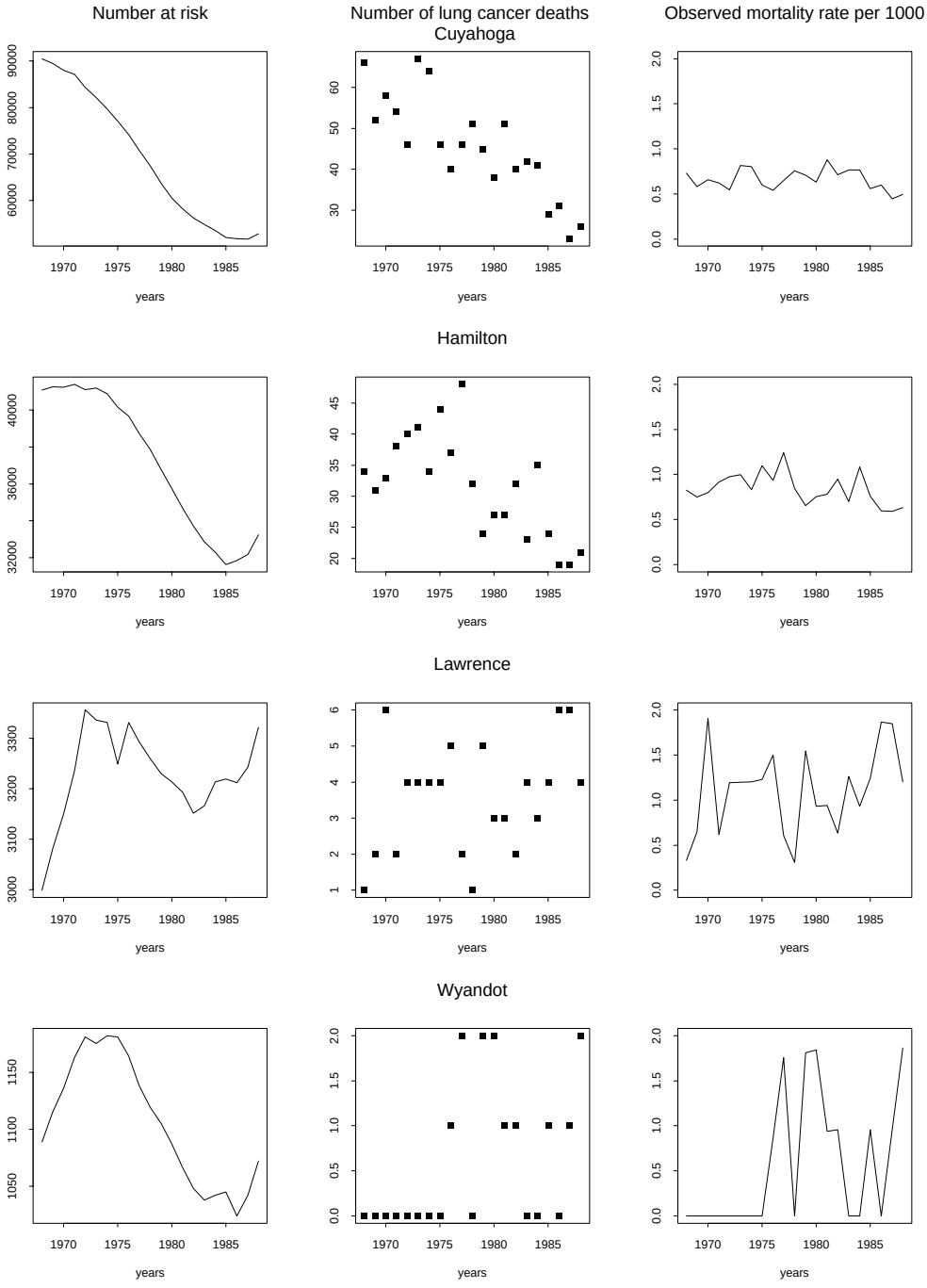

Figure 2: Temporal data for 45 to 54 year–old white males in four counties

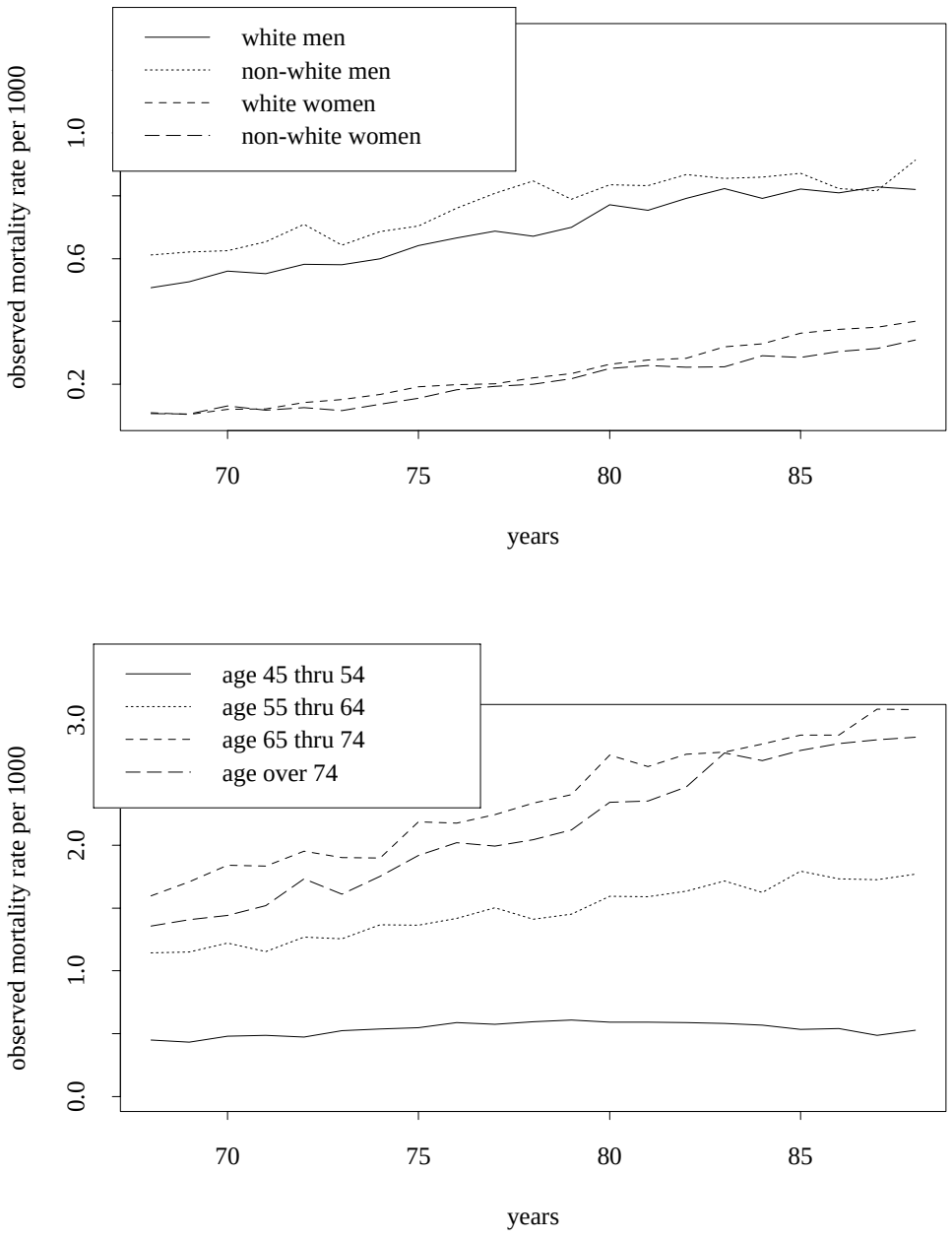

Figure 3: Temporal data, broken down by gender and race (top panel) and by agegroup (bottom panel), ignoring deaths below age 45.

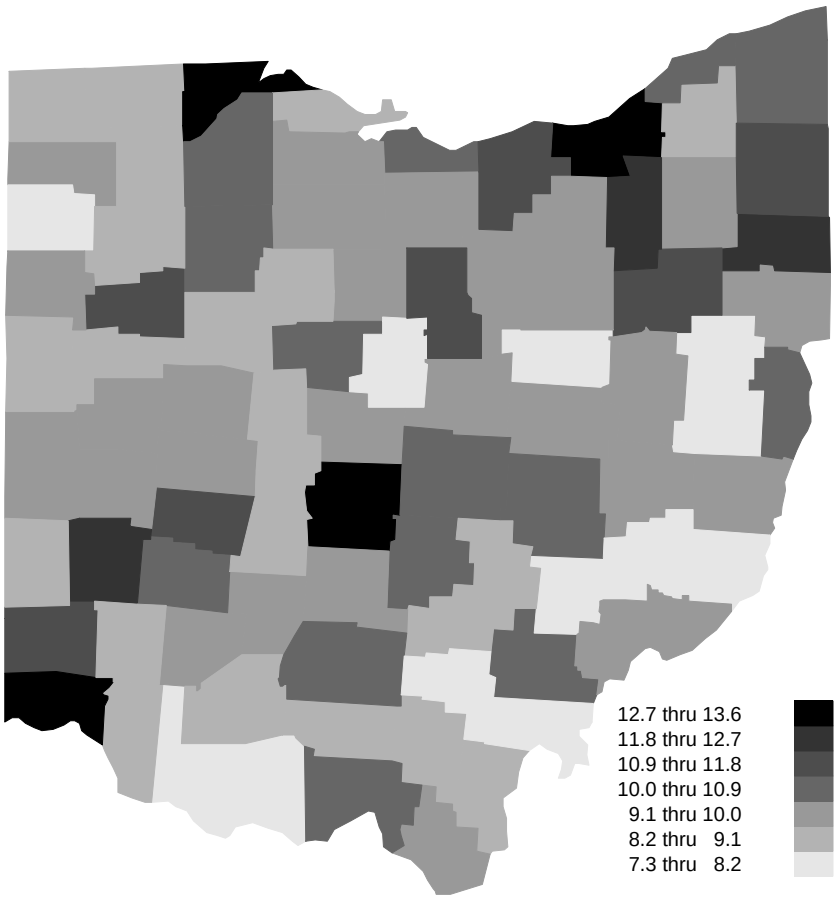

Figure 4: Kafadar–Tukey urbanization index for each county.

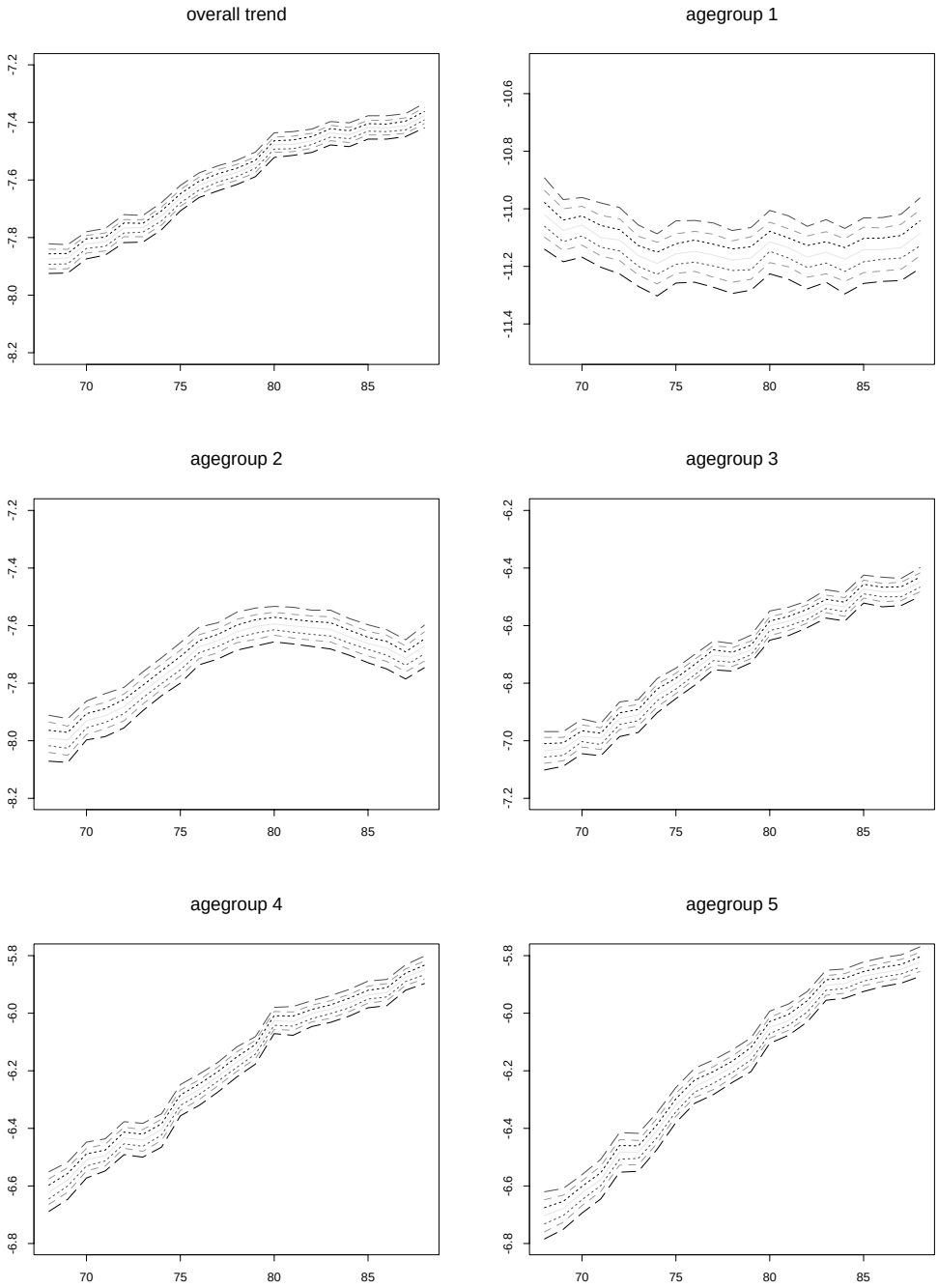

Figure 5: Medians and 50, 80 and 95% credible intervals for the overall time trend and for the aggregate of this with each of the five agegroup effects.

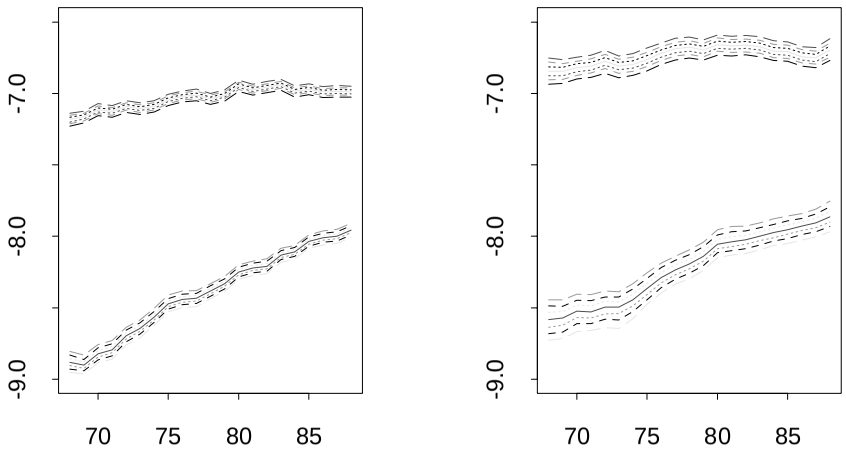

Figure 6: Medians and 50, 80 and 95% credible intervals for the effects due to gender $\times$ race, aggregated with the overall effect. Left panel: white men (above) and women. Right panel: non–white men (above) and women.

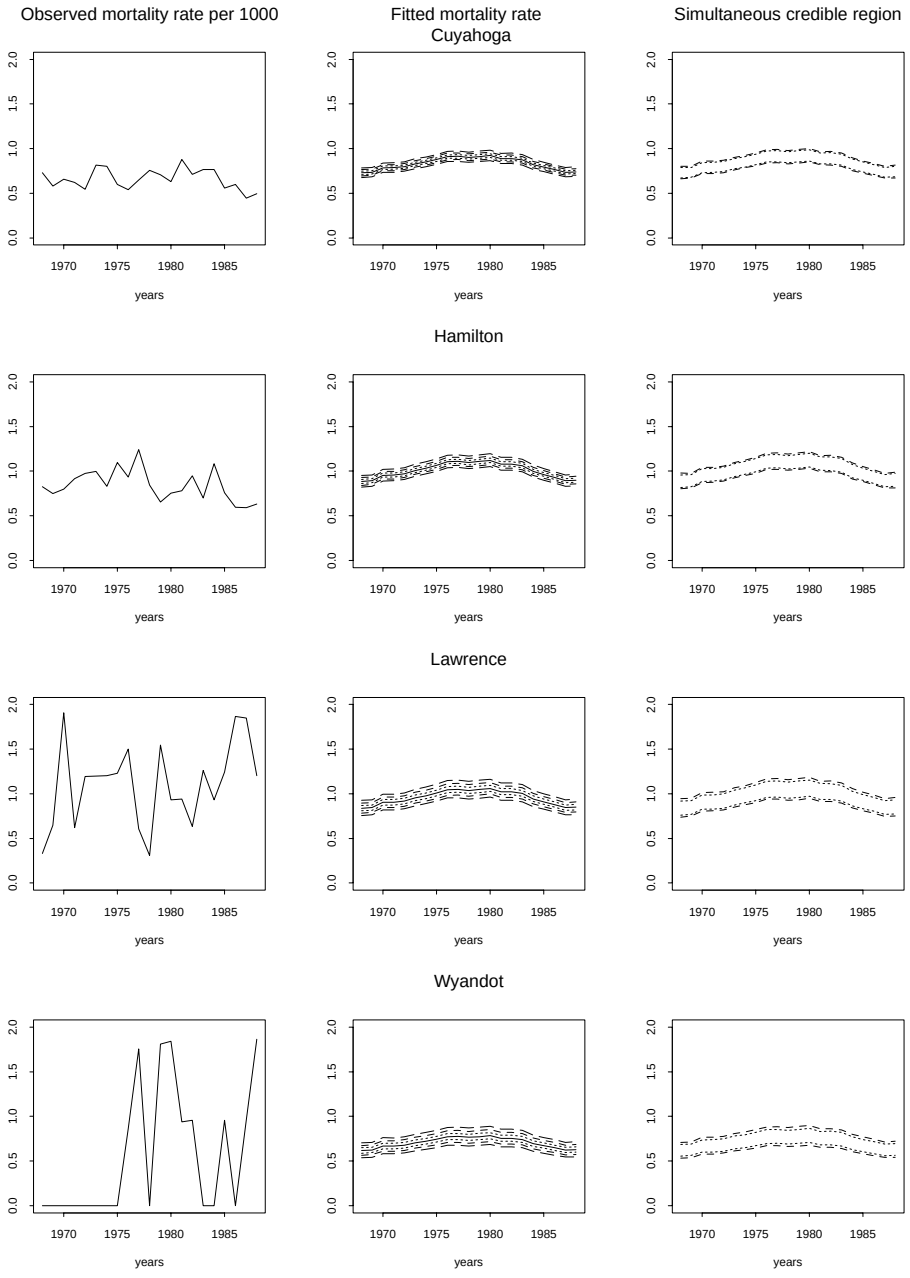

Figure 7: Observed and fitted mortality rates for 45 to 54 year–old white males in four counties. Centre panels: medians and 50, 80 and 90% credible intervals. Right–hand panels: 50 and 80% simultaneous credible regions.

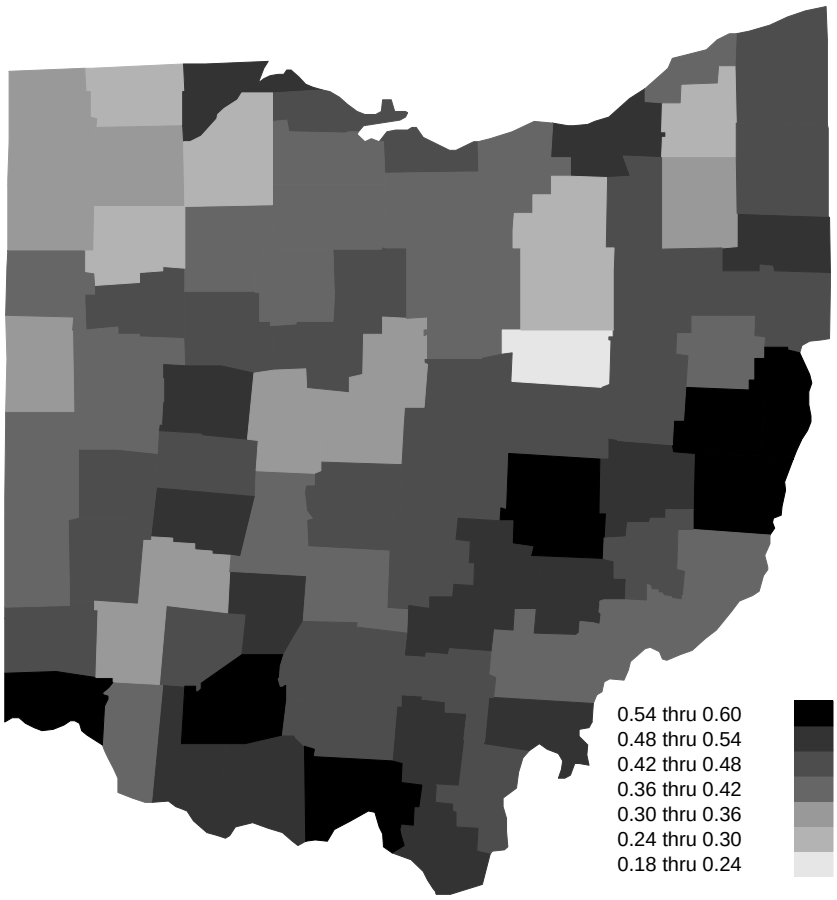

Figure 8: Overall fitted annual death rate $\times$ 1000 for each county.

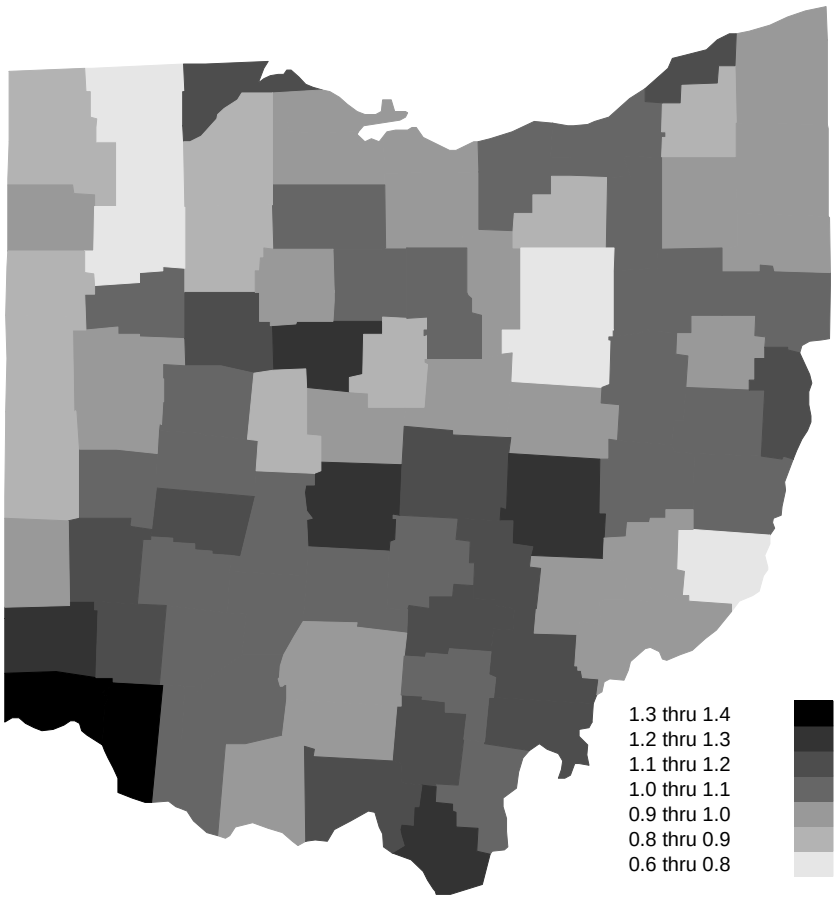

Figure 9: Posterior median for the adjusted relative risk in each county, without allowance for the Kafadar–Tukey index.

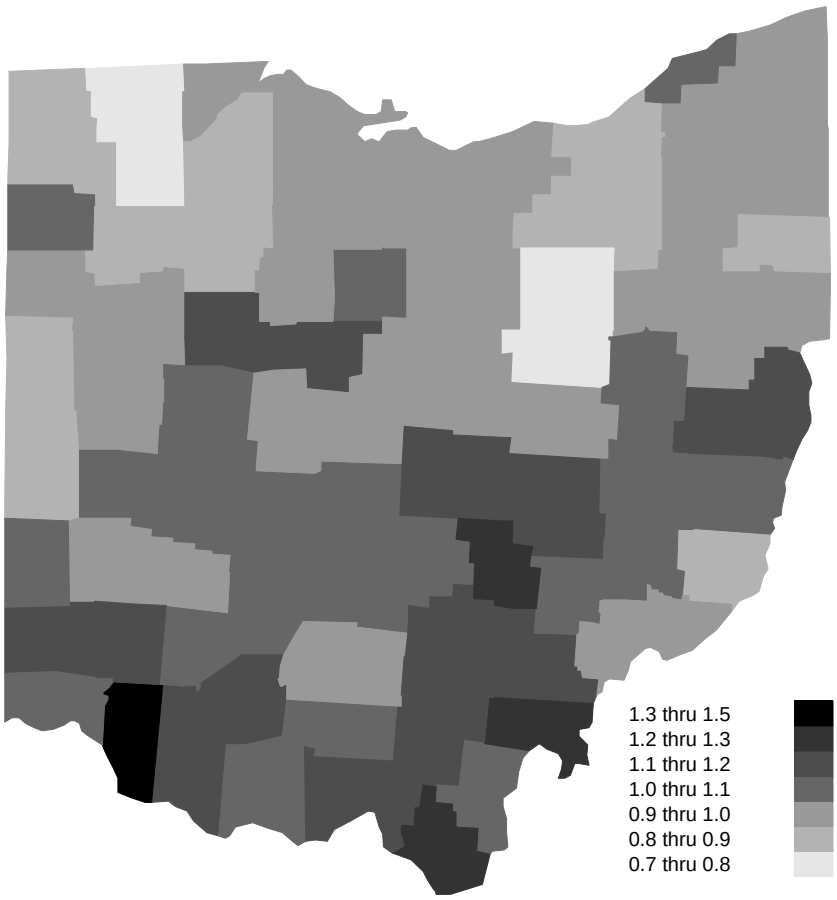

Figure 10: Posterior median for the adjusted relative risk in each county, allowing for the Kafadar–Tukey index.