BioMed Central

Methodology

Geostatistical analysis of disease data: accounting for spatial support and population density in the isopleth mapping of cancer mortality risk using area-to-point Poisson kriging

Pierre Goovaerts*

Address: BioMedware, Inc., Ann Arbor, MI, USA

Email: Pierre Goovaerts* - goovaerts@biomedware.com

* Corresponding author

Abstract

Background: Geostatistical techniques that account for spatially varying population sizes and spatial patterns in the filtering of choropleth maps of cancer mortality were recently developed. Their implementation was facilitated by the initial assumption that all geographical units are the same size and shape, which allowed the use of geographic centroids in semivariogram estimation and kriging. Another implicit assumption was that the population at risk is uniformly distributed within each unit. This paper presents a generalization of Poisson kriging whereby the size and shape of administrative units, as well as the population density, is incorporated into the filtering of noisy mortality rates and the creation of isopleth risk maps. An innovative procedure to infer the point-support semivariogram of the risk from aggregated rates (i.e. areal data) is also proposed.

Results: The novel methodology is applied to age-adjusted lung and cervix cancer mortality rates recorded for white females in two contrasted county geographies: 1) state of Indiana that consists of 92 counties of fairly similar size and shape, and 2) four states in the Western US (Arizona, California, Nevada and Utah) forming a set of 118 counties that are vastly different geographical units. Area-to-point (ATP) Poisson kriging produces risk surfaces that are less smooth than the maps created by a naïve point kriging of empirical Bayesian smoothed rates. The coherence constraint of ATP kriging also ensures that the population-weighted average of risk estimates within each geographical unit equals the areal data for this unit. Simulation studies showed that the new approach yields more accurate predictions and confidence intervals than point kriging of areal data where all counties are simply collapsed into their respective polygon centroids. Its benefit over point kriging increases as the county geography becomes more heterogeneous.

Conclusion: A major limitation of choropleth maps is the common biased visual perception that larger rural and sparsely populated areas are of greater importance. The approach presented in this paper allows the continuous mapping of mortality risk, while accounting locally for population density and areal data through the coherence constraint. This form of Poisson kriging will facilitate the analysis of relationships between health data and putative covariates that are typically measured over different spatial supports.

Background

Public health officials frequently use cancer mortality maps to identify areas of excess and their potential causes (e.g. environmental exposure or socio-demographic factors), as well as to guide surveillance and control activities. The interpretation and analysis of those choropleth maps faces three major hurdles: 1) the presence of extreme unreliable rates that typically occur for sparsely populated areas and/or less frequent cancers, 2) the visual bias resulting from the aggregation of health data within administrative units of widely different sizes and shapes, and 3) the mismatch of spatial supports for cancer rates and explanatory variables that prevents their direct use in correlation analysis and often implies an aggregation to the coarser geography. Common pitfalls include devoting unwarranted attention to a few oversized geographical units located in low population density areas or conducting regression analysis at scales that distort the true exposure/response relationship (i.e. ecological fallacy) [1]. What is needed is a coherent spatial methodology that allows both the filtering of the noise caused by the small number problem and the mapping of results as continuous surfaces without subjective administrative boundaries. Creation of isopleth risk maps from aggregated disease rates (i.e. areal data) has been the topic of several papers in the statistical and health science literatures. Geostatistical methods range from straightforward point kriging [2-5] to complex random field models that require distributional assumptions and computationally intensive parameter estimation using Markov Chain Monte Carlo (MCMC) techniques [6,7].

The simplest approach to create isopleth risk maps consists of assigning the aggregated rates to the geographic centroids of the administrative units, which are then interpolated to the nodes of a regular grid using point kriging. In these studies, the noise attached to rates estimated from small populations was either ignored [2], filtered using empirical Bayes smoothers prior to the interpolation [3], or incorporated in the interpolation procedure using binomial [4] or Poisson kriging [5]. When performing point kriging of areal data, the user makes the practical assumption that all the habitants of the administrative unit live at the same location and the measured rate thus refers to this specific location. This assumption is reasonable whenever the units of aggregation are small with respect to the spacing of the interpolation grid. For example, Oliver et al. [4] mapped the risk of childhood cancer over a 2 km grid in the West Midlands of England using binomial kriging and rates measured within small electoral wards. More recently, Ali et al. [5] created dysentery and cholera incidence maps using Poisson kriging and rates measured for patrilineally-related clusters of households. The assumption of point measurement support becomes clearly inappropriate when the administrative

units are counties or states [2,3], which calls for specific methods to incorporate the shape and size of those units in the analysis.

In theory, the system of kriging equations is flexible enough to incorporate data measured over a wide variety of supports and to accommodate either punctual or areal (classically termed "blocks" in the geostatistical literature) prediction supports [8,9]. Practical implementation of kriging in presence of irregular geographical units is, however, challenging mainly for the derivation of the areal covariance terms used in the kriging system [10]. Indeed, the covariance between two areas is approximated as the average of the point-support covariance computed between any two points discretizing these two areas. The first challenge is that the point-support semivariogram model cannot be inferred directly from the observations since only areal data are available. Instead, the semivariogram of areal data needs to be modelled, then deconvoluted to yield the point-support model. Second, the discretization of all the measurement and prediction supports, followed by the averaging of covariance values, can be CPU intensive. This partly explains why most geostatistical studies were restricted to the use of punctual data to conduct interpolation at punctual locations or over areas of regular size, such as blocks to be mined or environmental exposure units to remediate.

The geostatistical analysis of areal data has received increasing attention as kriging becomes more popular in the fields of remote sensing, social science, and medical geography. In particular, several authors [11-14] have implemented kriging of areal data to predict punctual or areal values, an approach referred to as "area-to-point" (ATP) or "area-to-area" (ATA) kriging following the terminology in Kyriakidis [14]. ATP kriging allows mapping the variability within geographical units while ensuring the coherence of the prediction so that, for example, disaggregated estimates of count data are non-negative and their sum is equal to the original aggregated count. Gotway and Young [12] applied ATA kriging to the mapping of the number of low birth weight (LBW) babies at the Census tract level, accounting for county-level LBW data and covariates measured over different spatial supports, such as a fine grid of ground-level PM_{10} concentrations or tract population.

Both ATA and ATP kriging have great potential for the analysis of cancer mortality maps, enabling the shape and size of administrative units to be incorporated into the smoothing of choropleth maps and the creation of isopleth risk maps, respectively. These methods must however account for the fact that disease rates are comprised of a numerator and a denominator, leading to data with varying degrees of reliability. Poisson kriging was recently

introduced to incorporate varying population sizes and spatial patterns in the processing of cancer mortality data [15,16]. Its implementation was facilitated by the initial assumption that all geographical units are the same size and shape, which allowed the use of geographic centroids in semivariogram estimation and kriging. Another implicit assumption was that the population at risk is uniformly distributed within each unit. Although these assumptions were reasonable for filtering cancer mortality maps that consist of geographical units of similar sizes, point Poisson kriging does not allow a rigorous treatment of the change of support, prohibiting for example the creation of continuous risk surfaces.

It is noteworthy that the increasingly popular full Bayesian modeling approach is overwhelmingly used with the conditional auto-regressive (CAR) model for defining the random effect associated with spatial autocorrelation [17-19]. This computationally convenient choice is reasonable if all geographical entities are of similar size and arranged in a regular pattern but it is not particularly attractive otherwise [6]. There have been several attempts to model the spatial correlation as a function of the distance between centroids of geographical units instead of using the arbitrary neighborhood relationship underlying the CAR model [20]. To quote Kelsall and Wakefield [6], "*This approach is quite simplistic, however, and again does not acknowledge the differing shapes and sizes and relative orientation of the areas*". The same authors describe one of the rare Bayesian models that accounts properly for the spatial support of the data, although it was implemented under the assumption of uniform population distribution within each area. Using semivariograms to model the spatial autocorrelation, they estimated the continuous underlying relative risk function for colorectal cancer mortality in 39 wards of the UK district of Birmingham.

This paper presents the first geostatistical study where the size and shape of administrative units, as well as the population density, is incorporated into the filtering of noisy mortality rates and the mapping of the corresponding risk at a fine scale (i.e. disaggregation). This generalization of the Poisson kriging algorithm introduced by Monestiez et al. [21,22] capitalizes on recent developments in the field of change of support and deconvolution of semivariogram estimated from irregular areal data [10]. The approach is illustrated using age-adjusted lung and cervix cancer mortality rates recorded for white females in two contrasted county geographies: 1) state of Indiana that consists of 92 counties of fairly similar size and shape, and 2) four states in the Western US (Arizona, California, Nevada and Utah) forming a set of 118 counties that are vastly different geographical units. Simulation studies are conducted to compare the performances of the new methodology to simple geostatistical approaches (i.e. point

kriging of raw or empirical Bayesian smoothed rates) for the same two county geographies. Performance criteria include the accuracy of the prediction of the underlying risk and the quantification of the attached uncertainty. The lack of realistic spatial models in a full Bayesian approach, coupled with the absence of user-friendly software to implement this type of model [23], prohibited the inclusion of this approach in the comparison study of isopleth mapping algorithms.

Methods
Data
The methodology to account for spatial support and population density in Poisson kriging will be illustrated using directly age-adjusted mortality rates for a frequent (lung) and less frequent (cervix) cancer. These data are part of the Atlas of Cancer Mortality in the United States [24] and were downloaded from http://www3.cancer.gov/atlas plus/download.html. The analysis focuses on white female rates recorded over the 1970–1994 period and adjusted using the 1970 population pyramid. Two areas with contrasted county geographies were considered: 1) state of Indiana (Region 1), and 2) four states in the Western US (Arizona, California, Nevada and Utah) that will be referred to as Region 2. The choice of these two specific geography areas was guided by the need to assess performances in two contrasted settings: 1) all geographical units have a fairly similar size and shape, which is the "ideal" situation for methods that ignore the spatial support of the data, and 2) geographical units display a wide range of sizes and shapes, which should favour any approach that implicitly accounts for the spatial support of the data in the analysis. The West coast provided a perfect example for the second type of geography (i.e. set of 118 vastly different counties), while Indiana includes a reasonable number (i.e. 92) of counties that are geometrically fairly similar.

Figure 1 (top graphs) shows the spatial distribution of age-adjusted mortality rates per 100,000 person-years for lung cancer in Region 1 and cervix cancer in Region 2. Following the recommendations of several studies on map color schemes [25,26], a double-ended color scheme with 10 equally-weighted classes (i.e. boundaries correspond to deciles of the histogram) was used: a gradient of red is used for rates higher than the median, while a gradient of blue is used for lower rates. The population-weighted average of mortality rates is 23.7 per 100,000 person-years for lung cancer and 2.85 per 100,000 person-years for cervix cancer. For Region 1, the population at risk in each county was computed as: 100,000 × the total number of "lung cancer" deaths over the 1970–1994 period divided by the corresponding age-adjusted mortality rate; both datasets are available on the website of the National Cancer Institute (NCI) of the US (URL as given above). A sim-

ilar procedure was conducted in Region 2, except that the most frequent breast cancer was used to avoid zero population estimates for a few sparsely populated counties where no cervix cancer mortality was reported during that period. The population sizes are mapped in Figure 1 (middle row) using a rainbow color scheme to avoid any confusion with the rate maps.

A visual inspection of the cervix cancer mortality map conveys the impression that rates are particularly high in the centre of the study area (Nye and Lincoln Counties), as well as in a few Northern California counties. This result must, however, be interpreted with caution since the population is not uniformly distributed across the study area and rates computed from sparsely populated counties tend to be less reliable. This effect, known as "small number problem" [27], is illustrated by the bottom scattergrams in Figure 1. In particular, cervix cancer mortality in excess of 5 deaths per 100,000 person-years is observed only for counties with a 25 year cumulated population smaller than 10^5. The use of administrative units to report the results (i.e. counties in this case) can also bias the interpretation: had the two counties with the highest rates been much smaller in size, these high values likely would have been perceived as less problematic. The higher frequency of lung cancer mortality, coupled with the narrower range of county sizes in Indiana, makes the interpretation of Region 1 cancer mortality map less hazardous. The highest rate is in fact reported for the most populated county (Marion) whose seat is Indianapolis.

To highlight the limitations of collapsing administrative units into their geographic centroids, the spatial distribution of population within each county is mapped at the top of Figure 2. These maps were produced by allocating the county-level population estimates of Figure 1 to a grid of 25 km² cells discretizing each study area, leading to 3,751 and 48,474 cells in Regions 1 and 2, respectively. The relative proportion of the county-level population within each 25 km² cell was retrieved from the readily available 2000 census block level data. These population maps illustrate the non-uniform repartition of population, in particular for the vast counties in Region 2. To account for the shape of geographical units and their heterogeneous population density, the distance between any two counties is here estimated as a population-weighted average of Euclidian distances between points discretizing the pair of counties:

$$Dist(v_\alpha, v_\beta) = \frac{1}{\sum\limits_{s=1}^{P_\alpha}\sum\limits_{s'=1}^{P_\beta} n(\mathbf{u}_s)n(\mathbf{u}_{s'})} \sum\limits_{s=1}^{P_\alpha}\sum\limits_{s'=1}^{P_\beta} n(\mathbf{u}_s)n(\mathbf{u}_{s'})\|\mathbf{u}_s - \mathbf{u}_{s'}\| \qquad (\text{Equation 1})$$

where P_α and P_β are the number of points $\mathbf{u}_s$ and $\mathbf{u}_{s'}$ used to discretize the two units or blocks v_α and v_β respectively.

The quantity n($\mathbf{u}_s$) represents the population size within the 25 km² cell centred on the discretizing point $\mathbf{u}_s$. For the examples of Figure 1, the discretizing points are identified with the nodes of the 5 km grid, yielding a total of 9 to 69 points per county in Indiana, and 11 to 2,082 discretizing points for the West Coast counties. The set of block-to-block distances (Equation (1)) are plotted against Euclidian distances between polygon centroids depicted by red dots in Figure 2 (middle maps). The scatterplots at the bottom of Figure 2 indicate a high correlation between the two sets of distances; discrepancies essentially occur for neighboring counties (i.e. for small distances). The correlation is slightly smaller for Region 2 where county shapes and sizes vary greatly. Block-to-block distances are numerically very close to the distances computed between population-weighted centroids depicted by blue triangles in Figure 2 (middle maps).

Poisson kriging: the centroid-based approach

The geostatistical methodology for the estimation of risk values from empirical frequencies, and its performance relative to common smoothers, is described in details in Goovaerts [15]. This section provides a brief recall of the centroid-based implementation of Poisson kriging (PK) for prediction of aggregated risk values. The approach is then generalized to account for spatial supports and population density in both the case of area-to-area and area-to-point predictions.

For a given number N of geographical units v_α (e.g. counties), denote the observed mortality rates (areal data) as $z(v_\alpha) = d(v_\alpha)/n(v_\alpha)$, where $d(v_\alpha)$ is the number of recorded mortality cases and $n(v_\alpha)$ is the size of the population at risk. Let us assume for now that all units v_α have similar shapes and sizes, with a uniform population density. A straightforward approach to implement geostatistics is to assume that each unit or measurement support is a single point. This simplification amounts at assigning each measured rate to the geographic centroid of the unit over which it has been recorded. For example, let $\mathbf{u}_\alpha = (x_\alpha, y_\alpha)$ be the vector of spatial coordinates for the centroid of the unit v_α. The disease count $d(\mathbf{u}_\alpha)$ is interpreted as a realization of a random variable $D(\mathbf{u}_\alpha)$ that follows a Poisson distribution with one parameter (expected number of counts) that is the product of the population size $n(\mathbf{u}_\alpha)$ by the local risk $R(\mathbf{u}_\alpha)$.

The risk over a given unit v_α is estimated as a linear combination of the rate observed for that unit, $z(\mathbf{u}_\alpha)$, and in (K-1) neighboring units:

$$\hat{r}_{PK}(\mathbf{u}_\alpha) = \sum_{i=1}^{K} \lambda_i(\mathbf{u}_\alpha) z(\mathbf{u}_i) \qquad (\text{Equation 2})$$
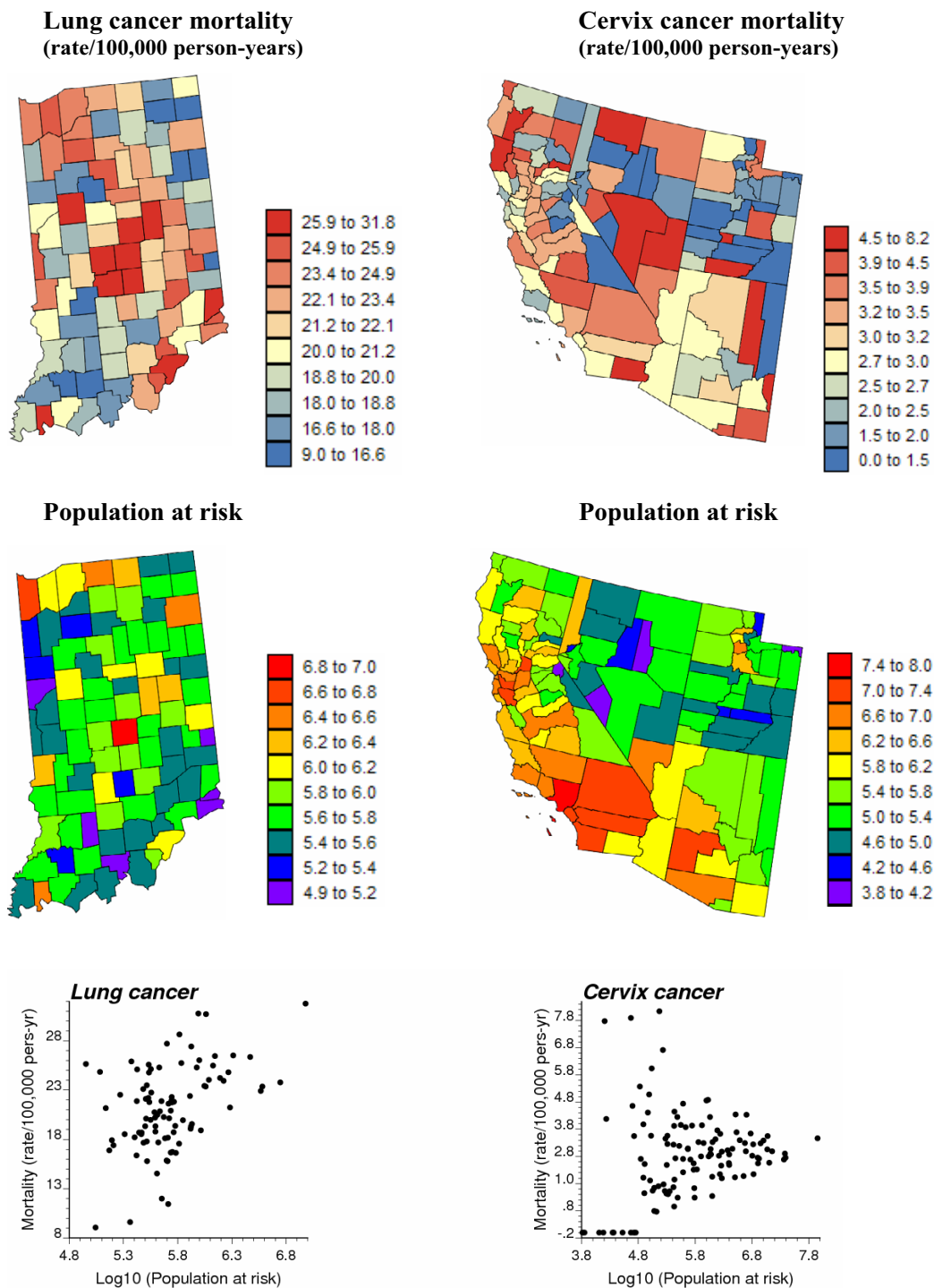
Figure 1
Geographic distribution of age-adjusted lung and cervix cancer mortality rates and the populations at risk. For the two top maps, the fill color in each county represents the age-adjusted mortality rates per 100,000 person-years recorded over the period 1970–1994 for white females (class boundaries correspond to deciles of the histogram of rates). The middle maps represent the population at risk which was back-calculated from the rate and count data (lognormal scale). The bottom scatterplots illustrate the larger variance of rates computed from sparsely populated counties, in particular for the least frequent cervix cancer.
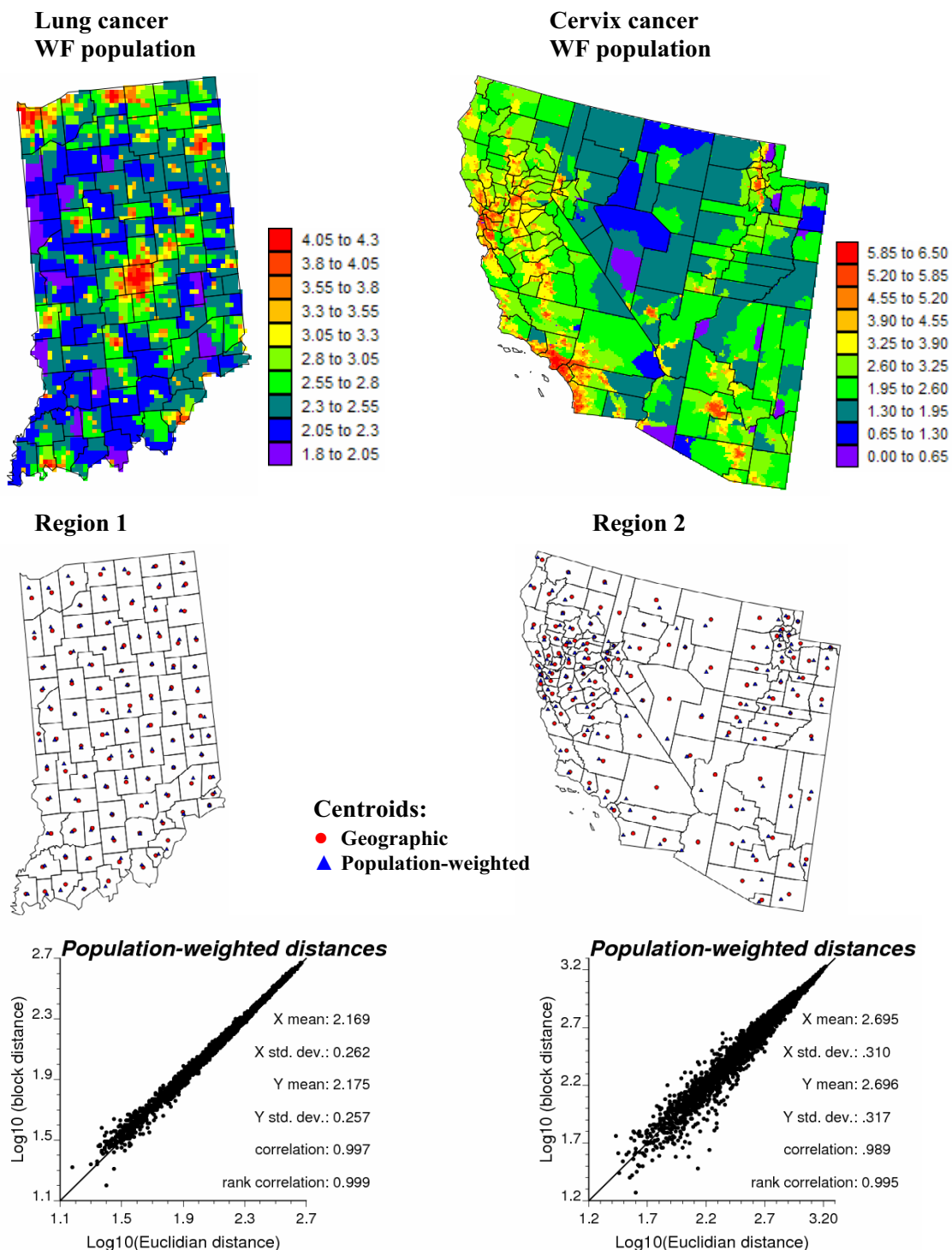
Figure 2
High-resolution population maps and their use to compute population-weighted distances between counties.
The county-level population estimates of Figure 1 were allocated to a grid of 25 km² cells according to the 2000 census block level data. The scattergrams plot Euclidian distances between county geographic centroids, depicted by red dots in the middle maps, versus a "block distance" that accounts for the shape of counties and the distribution of the population (Equation 1). Population-weighted centroids are depicted by blue triangles in the middle maps.

where $\lambda_i(\mathbf{u}_\alpha)$ is the weight assigned to the rate $z(\mathbf{u}_i)$ when estimating the risk at $\mathbf{u}_\alpha$. The K weights are the solution of the following system of linear equations:

$$\sum_{j=1}^{K} \lambda_j(\mathbf{u}_\alpha)\left[C_R(\mathbf{u}_i - \mathbf{u}_j) + \delta_{ij}\frac{m^*}{n(\mathbf{u}_i)} \right] + \mu(\mathbf{u}_\alpha) = C_R(\mathbf{u}_i - \mathbf{u}_\alpha) \quad i = 1,...,K$$
$$\sum_{j=1}^{K} \lambda_j(\mathbf{u}_\alpha) = 1 \qquad\qquad (\text{Equation } 3)$$

where $\delta_{ij} = 1$ if $\mathbf{u}_i = \mathbf{u}_j$ and 0 otherwise, and m^* is the population-weighted mean of the N rates. The "error variance" term, $m^*/n(\mathbf{u}_i)$, leads to smaller weights for less reliable data (i.e. rates measured over smaller populations). The prediction variance associated with the estimate (2) is computed as:

$$\sigma_{PK}^2(\mathbf{u}_\alpha) = C_R(0) - \sum_{i=1}^{K} \lambda_i(\mathbf{u}_\alpha) C_R(\mathbf{u}_i - \mathbf{u}_\alpha) - \mu(\mathbf{u}_\alpha) \qquad (\text{Equation } 4)$$

To solve the kriging system (Equation 3), one needs a model for the covariance of the risk, $C_R(\mathbf{h})$, or equivalently its semivariogram $\gamma_R(\mathbf{h}) = C_R(0) - C_R(\mathbf{h})$. Following Monestiez et al. [21,22] the semivariogram of the risk is estimated as:

$$\hat{\gamma}_R(\mathbf{h}) = \frac{1}{2\sum_{\alpha=1}^{N(\mathbf{h})} \frac{n(\mathbf{u}_\alpha)n(\mathbf{u}_\alpha + \mathbf{h})}{n(\mathbf{u}_\alpha) + n(\mathbf{u}_\alpha + \mathbf{h})}} \sum_{\alpha=1}^{N(\mathbf{h})} \left\{ \frac{n(\mathbf{u}_\alpha)n(\mathbf{u}_\alpha + \mathbf{h})}{n(\mathbf{u}_\alpha) + n(\mathbf{u}_\alpha + \mathbf{h})}\left[z(\mathbf{u}_\alpha) - z(\mathbf{u}_\alpha + \mathbf{h}) \right]^2 - m^* \right\} \qquad (\text{Equation } 5)$$

The different spatial increments $[z(\mathbf{u}_\alpha) - z(\mathbf{u}_\alpha + \mathbf{h})]^2$ are weighted by a function of their respective population sizes, $n(\mathbf{u}_\alpha)n(\mathbf{u}_\alpha + \mathbf{h})/(n(\mathbf{u}_\alpha) + n(\mathbf{u}_\alpha + \mathbf{h}))$, a term which is inversely proportional to their standard deviation. More importance is thus given to the more reliable data pairs (i.e. smaller standard deviations).

Area-to-Area (ATA) Poisson kriging
In presence of geographical units with very different shapes and sizes, it is overly simplistic to assimilate each unit v_α to its geographic centroid $\mathbf{u}_\alpha$. The spatial support of each unit needs to be accounted for in both the semivariogram inference and kriging. Following the terminology in Kyriakidis [14], ATA kriging refers to the case where both the prediction and measurement supports are areas instead of points. The PK estimate (2) for the areal risk value $r(v_\alpha)$ thus becomes:

$$\hat{r}_{PK}(v_\alpha) = \sum_{i=1}^{K} \lambda_i(v_\alpha) z(v_i) \qquad (\text{Equation } 6)$$

The weights $\lambda_i(v_\alpha)$ are computed by solving the following kriging system:

$$\sum_{j=1}^{K} \lambda_j(v_\alpha)\left[\bar{C}_R(v_i, v_j) + \delta_{ij}\frac{m^*}{n(v_i)} \right] + \mu(v_\alpha) = \bar{C}_R(v_i, v_\alpha) \quad i = 1,...,K$$
$$\sum_{j=1}^{K} \lambda_j(v_\alpha) = 1. \qquad\qquad (\text{Equation } 7)$$

The main difference between systems (3) and (7) is that point-to-point covariance terms $C_R(\mathbf{u}_i - \mathbf{u}_j)$ are replaced by area-to-area covariances $\bar{C}_R(v_i, v_j) = \mathrm{Cov}\{Z(v_i), Z(v_j)\}$. Those covariances are numerically approximated by averaging the point-support covariance $C(\mathbf{h})$ computed between any two locations discretizing the areas v_i and v_j:

$$\bar{C}_R(v_i, v_j) = \frac{1}{\sum_{s=1}^{P_i}\sum_{s'=1}^{P_j} w_{ss'}} \sum_{s=1}^{P_i}\sum_{s'=1}^{P_j} w_{ss'} C(\mathbf{u}_s, \mathbf{u}_{s'}) \qquad (\text{Equation } 8)$$

where P_i and P_j are the number of points used to discretize the two areas v_i and v_j, respectively. The weights $w_{ss'}$ are computed as the product of population sizes within the square cells centred on the discretizing point $\mathbf{u}_s$ and $\mathbf{u}_{s'}$:

$$w_{ss'} = n(\mathbf{u}_s) \times n(\mathbf{u}_{s'}) \text{ with } \sum_{s=1}^{P_i} n(\mathbf{u}_s) = n(v_i) \text{ and } \sum_{s'=1}^{P_j} n(\mathbf{u}_{s'}) = n(v_j) \qquad (\text{Equation } 9)$$

The kriging variance for the ATA kriging estimator is computed as:

$$\sigma_{PK}^2(v_\alpha) = \bar{C}_R(v_\alpha, v_\alpha) - \sum_{i=1}^{K} \lambda_i(v_\alpha)\bar{C}_R(v_i, v_\alpha) - \mu(v_\alpha) \qquad (\text{Equation } 10)$$

where $\bar{C}_R(v_\alpha, v_\alpha)$ is the within-area covariance that is computed according to Equation (8) with $v_i = v_j = v_\alpha$.

Area-to-Point (ATP) Poisson kriging
A particular case of ATA kriging is when the prediction support is so small that it can be assimilated to a point $\mathbf{u}_s$, leading to the following area-to-point kriging estimator and kriging variance:

$$\hat{r}_{PK}(\mathbf{u}_s) = \sum_{i=1}^{K} \lambda_i(\mathbf{u}_s) z(v_i) \qquad (\text{Equation } 11)$$

$$\sigma_{PK}^2(\mathbf{u}_s) = C_R(0) - \sum_{i=1}^{K} \lambda_i(\mathbf{u}_s)\bar{C}_R(v_i, \mathbf{u}_s) - \mu(\mathbf{u}_s) \qquad (\text{Equation } 12)$$

The kriging weights and the Lagrange parameter $\mu(\mathbf{u}_s)$ are computed by solving the following system of linear equations:

$$\sum_{j=1}^{K} \lambda_j(\mathbf{u}_s)\left[\bar{C}_R(v_i, v_j) + \delta_{ij}\frac{m^*}{n(v_i)} \right] + \mu(\mathbf{u}_s) = \bar{C}_R(v_i, \mathbf{u}_s) \quad i = 1,...,K$$
$$\sum_{j=1}^{K} \lambda_j(\mathbf{u}_s) = 1. \qquad\qquad (\text{Equation } 13)$$

The ATP kriging system is similar to the ATA kriging system (Equation 7), except for the right-hand-side term where the area-to-area covariances $\bar{C}_R(v_i, v_\alpha)$ are replaced

by area-to-point covariances $\bar{C}_R(v_i, \mathbf{u}_s)$ that are approximated as:

$$\bar{C}_R(v_i, \mathbf{u}_s) = \frac{1}{\sum\limits_{s'=1}^{P_i} w_{s's}} \sum_{s'=1}^{P_i} w_{s's} C(\mathbf{u}_{s'}, \mathbf{u}_s) \qquad (\text{Equation 14})$$

where P_i is the number of points used to discretize the area v_i and the weights $w_{s's}$ are computed according to expression (9).

ATP kriging can be conducted at each node of a grid covering the study area, resulting in a continuous (isopleth) map of mortality risk and reducing the visual bias that is typically associated with the interpretation of choropleth maps. Another interesting property of the ATP kriging estimator is its coherence: the population-weighted average of the risk values estimated at the P_α points $\mathbf{u}_s$ discretizing a given entity v_α yields the ATA risk estimate for this entity:

$$\hat{r}_{PK}(v_\alpha) = \frac{1}{n(v_\alpha)} \sum_{s=1}^{P_\alpha} n(\mathbf{u}_s) \hat{r}_{PK}(\mathbf{u}_s) \qquad (\text{Equation 15})$$

Constraint (15) is satisfied if the same K areal data are used for the ATP kriging of the P_α risk values. Indeed, in this case the population-weighted average of the right-hand-side covariance terms of the K ATP kriging systems is equal to the right-hand-side covariance of the single ATA kriging system:

$$\frac{1}{n(v_\alpha)}\sum_{s=1}^{P_\alpha} n(\mathbf{u}_s)\bar{C}_R(v_i,\mathbf{u}_s) = \frac{1}{n(v_\alpha)}\sum_{s=1}^{P_\alpha} n(\mathbf{u}_s)\left[\frac{1}{n(v_i)}\sum_{s'=1}^{P_i} n(\mathbf{u}_{s'})C(\mathbf{u}_{s'},\mathbf{u}_s)\right]$$
$$= \bar{C}_R(v_i, v_\alpha), \qquad (\text{Equation 16})$$

per relations (8), (9) and (14). Therefore, the following relationship exists between the two sets of ATA and ATP kriging weights:

$$\lambda_i(v_\alpha) = \frac{1}{n(v_\alpha)} \sum_{s=1}^{P_\alpha} n(\mathbf{u}_s) \lambda_i(\mathbf{u}_s) \quad i = 1,...,K \qquad (\text{Equation 17})$$

Deconvolution of the semivariogram of the risk
To solve both ATA and ATP kriging systems, one needs to know the point-support covariance of the risk $C(\mathbf{h})$, or equivalently the point-support semivariogram $\gamma(\mathbf{h})$. This function cannot be estimated directly from the observed rates, since only aggregated data are available. Derivation of a point-support semivariogram from the "regularized" experimental semivariogram computed from areal data is called "deconvolution" [14,28]. Unlike the mining applications that have led to the initial development of geostatistical deconvolution [8], the size and shape of geographical units in health science applications can vary greatly, which makes the deconvolution much more challenging. Goovaerts [10] developed an innovative

approach to conduct the deconvolution in presence of irregular units and heterogeneous population distributions. Like most inverse problems, the deconvolution is best tackled using an iterative procedure. The basic idea is to seek the point-support model that, once regularized, is the closest to the model fitted to areal data. Only the most salient features of the deconvolution procedure are described hereafter, and the reader is referred to Goovaerts [10] for a detailed presentation of the algorithm and demonstration of its performances in simulation studies.

The deconvolution is based on the following relationship between the regularized semivariogram model $\gamma_v(h)$ fitted to areal data and area-to-area semivariogram values that are a function of the unknown point-support model $\gamma(\mathbf{h})$:

$$\gamma_v(h) = \bar{\gamma}(v, v_h) - \bar{\gamma}_h(v, v) \qquad (\text{Equation 18})$$

The area-to-area semivariogram value $\bar{\gamma}(v, v_h)$ represents the mean value of the point-support semivariogram between an arbitrary point in the support v and another in the translated support v_h. Because each area (e.g. county) has potentially a different size and shape, one needs to consider any two areas that can be encountered across the study area and average the semivariogram values over all $N(h)$ pairs of areas separated by the distance h. More precisely, the area-to-area semivariogram value is estimated as:

$$\bar{\gamma}(v, v_h) = \frac{1}{N(h)} \sum_{\alpha=1}^{N(h)} \bar{\gamma}(v_\alpha, v_{\alpha+h}) \qquad (\text{Equation 19})$$

The semivariogram value between any two areas, v_α and $v_{\alpha+h}$, separated by the population-based distance h (Equation 1) is computed as:

$$\bar{\gamma}(v_\alpha, v_{\alpha+h}) = \frac{1}{P_\alpha P_{\alpha+h}} \sum_{s=1}^{P_\alpha} \sum_{s'=1}^{P_{\alpha+h}} \gamma(\mathbf{u}_s, \mathbf{u}_{s'}) \qquad (\text{Equation 20})$$

where P_α and $P_{\alpha+h}$ are the number of points used to discretize the two areas v_α and $v_{\alpha+h}$, respectively.

The second term in the regularization expression (Equation 18) is the within-area semivariogram value $\bar{\gamma}_h(v, v)$. It fluctuates as a function of the separation distance h since the size of the areas used in semivariogram computation varies as a function of the distance between them: smaller areas tend to be paired at shorter distances, while larger areas can only be paired for distances that exceed half their minimum dimension. This quantity is estimated as the arithmetical average of within-area semivariogram values for any pair of areas separated by a given distance h:

$$\overline{\gamma}_h(v,v) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [\overline{\gamma}(v_\alpha, v_\alpha) + \overline{\gamma}(v_{\alpha+h}, v_{\alpha+h})] \qquad (\text{Equation 21})$$

where $\overline{\gamma}(v_\alpha, v_\alpha)$ and $\overline{\gamma}(v_{\alpha+h}, v_{\alpha+h})$ are computed using Equation (20) with h = 0.

The deconvolution procedure starts with the choice of an initial point-support model $\gamma^{(0)}(\mathbf{h})$; for example the model $\gamma_R(\mathbf{h})$ fitted to the areal data. This model is regularized according to expression (18), and the theoretically regularized model, $\gamma_v^{(0)}(h)$, is compared to the model inferred from areal data, $\gamma_R(\mathbf{h})$. The optimization criterion is the relative difference between the two curves, denoted D. A new candidate point-support semivariogram $\gamma^{(1)}(\mathbf{h})$ is derived by rescaling of the initial point-support model $\gamma^{(0)}(\mathbf{h})$, and then regularized according to expression (18). Model $\gamma^{(1)}(\mathbf{h})$ becomes the new optimum if it lowers the deviation between the theoretically regularized semivariogram model $\gamma_v^{(1)}(h)$ and the model fitted to areal data, that is if $D^{(1)} < D^{(0)}$. Rescaling coefficients are then updated to account for the difference between $\gamma_v^{(1)}(h)$ and $\gamma_R(\mathbf{h})$, leading to a new candidate model $\gamma^{(2)}(\mathbf{h})$ for the next iteration. The procedure stops when the maximum number of allowed iterations has been tried (e.g. 35 in this paper) or the decrease in the D statistic becomes negligible from one iteration to the next. The use of lag-specific rescaling coefficients provides enough flexibility to modify the initial shape of the point-support semivariogram and makes the deconvolution insensitive to the initial solution adopted.

Simple approaches to perform area-to-point kriging

A couple of geostatistical approaches have been proposed in the health science literature to derive continuous maps of health outcomes from areal data. The most straightforward method is to interpolate the raw rates to the nodes of a regular grid using point ordinary kriging [2]:

$$\hat{z}_{OK}(\mathbf{u}_s) = \sum_{i=1}^{K} \lambda_i(\mathbf{u}_s) z(\mathbf{u}_i) \qquad (\text{Equation 22})$$

The corresponding kriging variance is computed as:

$$\sigma_{OK}^2(\mathbf{u}_s) = C(0) - \sum_{i=1}^{K} \lambda_i(\mathbf{u}_s) C(\mathbf{u}_i - \mathbf{u}_s) - \mu(\mathbf{u}_s) \qquad (\text{Equation 23})$$

The weights $\lambda_i(\mathbf{u}_s)$ are the solution of the following ordinary kriging system:

$$\sum_{j=1}^{K} \lambda_j(\mathbf{u}_s) C(\mathbf{u}_i - \mathbf{u}_j) + \mu(\mathbf{u}_s) = C(\mathbf{u}_i - \mathbf{u}_s) \quad i = 1,...,K$$
$$\sum_{j=1}^{K} \lambda_j(\mathbf{u}_s) = 1 \qquad (\text{Equation 24})$$

The covariance C(**h**) is derived from the model fitted to the experimental semivariogram of raw rates that is calculated as:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} [z(\mathbf{u}_\alpha) - z(\mathbf{u}_\alpha + \mathbf{h})]^2 \qquad (\text{Equation 25})$$

This approach overlooks two critical facts: 1) rates can be very unreliable if recorded over sparsely populated geographical units, and 2) the spatial support of the data (i.e. county) can not be simply equalled to the prediction support (i.e. grid node). In other words, it is unrealistic to collapse vastly different administrative units into their geographic centroids. Another limitation of this approach is its lack of coherence: there is no guarantee that the average of kriging estimates for all P_α points discretizing the county v_α yields the areal data $z(v_\alpha)$:

$$z(v_\alpha) \neq \frac{1}{P_\alpha} \sum_{s=1}^{P_\alpha} z_{OK}(\mathbf{u}_s) \qquad (\text{Equation 26})$$

To correct for the "small number problem", Berke [3] proposed to replace the raw rates $z(\mathbf{u}_i)$ in Equations (22) and (25) by rates $\hat{r}_{GBS}(\mathbf{u}_i)$ obtained using global empirical Bayes smoothers. Each smoothed rate is easily computed as a linear combination of the raw rate $z(\mathbf{u}_i)$ and the global mean m*:

$$\hat{r}_{GBS}(\mathbf{u}_i) = \lambda(\mathbf{u}_i) z(\mathbf{u}_i) + [1 - \lambda(\mathbf{u}_i)]m^* \quad i = 1,...,N \quad (\text{Equation 27})$$

The Bayes shrinkage factor $\lambda(\mathbf{u}_i)$ is computed as:

$$\lambda(\mathbf{u}_i) = \begin{cases} \dfrac{s^2 - m^*/\overline{n}}{s^2 - m^*/\overline{n} + m^*/n(\mathbf{u}_i)} & \text{if } s^2 \geq m^*/\overline{n} \\ 0 & \text{otherwise} \end{cases} \qquad (\text{Equation 28})$$

where m* and s^2 are the population-weighted sample mean and variance of rates, and $\overline{n}$ is the average population size across the study area. Whenever the rate $z(\mathbf{u}_i)$ is based on small population sizes $n(\mathbf{u}_i)$ relative to the average size $\overline{n}$, the factor $\lambda(\mathbf{u}_i)$ is small and the Bayesian estimate (Equation 27) is close to the global mean m*.

The kriging of empirical Bayesian smoothed rates suffers from the same shortcomings as the kriging of raw rates:

lack of coherence and ignorance of the spatial support of the data. To attenuate the smoothing effect caused by the use of a global mean in the Bayes smoother (Equation 27), in this paper kriging was also applied to local Bayes estimates. Local Bayes smoothers are calculated similarly except that all the statistics (i.e. the population-weighted sample mean and variance, population size) are computed within local search windows [29]; for example using the K neighboring observed rates. The estimator thus becomes:

$$\hat{r}_{LBS}(\mathbf{u}_i) = \lambda(\mathbf{u}_i)z(\mathbf{u}_i) + [1 - \lambda(\mathbf{u}_i)]m^*(\mathbf{u}_i) \quad \text{(Equation 29)}$$

where $m^*(\mathbf{u}_i)$ is the population-weighted average of the rates within the search window. The Bayes shrinkage factor $\lambda(\mathbf{u}_i)$ is now computed as:

$$\lambda(\mathbf{u}_i) = \begin{cases} \dfrac{s^2(\mathbf{u}_i) - m^*(\mathbf{u}_i)/\bar{n}(\mathbf{u}_i)}{s^2(\mathbf{u}_i) - m^*(\mathbf{u}_i)/\bar{n}(\mathbf{u}_i) + m^*(\mathbf{u}_i)/n(\mathbf{u}_i)} & \text{if } s^2(\mathbf{u}_i) \geq m^*(\mathbf{u}_i)/\bar{n}(\mathbf{u}_i) \\ 0 & \text{otherwise} \end{cases}$$

As for global Bayes smoothers, the relative weight $\lambda(\mathbf{u}_i)$ assigned to the observed rate $z(\mathbf{u}_i)$ is smaller for less densely populated counties. For counties with similar population sizes, the factor $\lambda(\mathbf{u}_i)$ is also smaller in regions of greater homogeneity, as measured by a lower local variance $s^2(\mathbf{u}_i)$.

Results and discussion
Analysis of lung and cervix cancer data
County-level estimates of mortality risk
Figure 3 (top graphs, red curve) shows the omnidirectional semivariograms of lung and cervix cancer mortality risk computed from county-level rates using estimator (5) and the distance measure (1). The experimental semivariogram was fitted using a spherical model with a range of 113 km for lung cancer in Indiana and 437 km for cervix cancer in Region 2. Each model was deconvoluted using the iterative procedure and the high-resolution population maps displayed in Figure 2. For both regions, the procedure stopped once a small (i.e. <1%) decrease in the D statistic occurred three times, after 28 iterations for Region 1 and 10 iterations for Region 2. As expected, the point-support semivariogram model (green curve) has a higher sill since the punctual process has a larger variance than its aggregated form. The difference between the sills is particularly large for Region 1, which confirms the stronger impact of regularization for processes with smaller range of autocorrelation [[30], p. 465]. Applying expression (18) to the point-support model yields a theoretically regularized semivariogram model (blue curve) that is close to the one fitted to experimental values, which validates the consistency of the deconvolution.

The deconvoluted semivariogram models were used to estimate aggregated risk values at the county level in both regions (ATA kriging); see Figure 4 (bottom maps). Figure 4 also shows the maps produced by global and local empirical Bayes smoothers. In all cases, the estimation was based on $K = 32$ closest observations which were selected according to the population-weighted distance between counties for ATA kriging and the Euclidian distance between the county population-weighted centroids for the two other methods. All maps are smoother than the map of raw rates since the noise due to small population sizes is filtered. The smoothing is particularly strong for the less frequent cervix cancer with a one order of magnitude drop in the variance; see Table 1 (top rows). The variance of smoothed mortality rates is half the variance of raw rates for lung cancer. Table 1 indicates that for both cancers the smoothing is slightly smaller for ATA kriging versus local Bayes smoothers. The global EBS estimates are the least variable, in particular for cervix cancer.

The three risk maps of Region 1 are relatively similar; see Figure 4 (left column). All methods smoothed low lung cancer rates recorded in a few North-western and North-eastern counties characterized by smaller population sizes. For example, after application of Poisson kriging the minimum rate increased from 9.071 to 13.22 deaths/100,000 habitants; see Table 1. Since the highest lung cancer mortality rates are observed in the most populated counties in Indiana, the maximum rates remain practically the same after smoothing.

Differences between methods are much more pronounced for cervix cancer mortality in Region 2; see Figure 4 (right column). Unlike the case of lung cancer, both high and low cancer rates are smoothed. Except on the local EBS map, the high risk area formed by two central counties in Figure 1 faded, which illustrates the potential pitfalls associated with the interpretation of the map of observed rates. The highest risk (4.081 deaths/100,000 habitants for ATA kriging) is predicted for Kern County, just west of Santa Barbara County (California). Zero cervix cancer mortality rates recorded in sparsely populated counties in Utah were also smoothed, leading to minimum values of 1.92–2.28 deaths/100,000 habitants. Global empirical Bayes method causes a strong smoothing of the raw rates, in particular in Utah where outside Salt Lake City most county rates shrunk towards the global mean of 2.85 per 100,000 person-years. The two other methods (ATA kriging and local EBS) create risk maps that show a large contrast between a low risk cluster in Utah and a high risk cluster in Southern California and Southern Nevada. Still, these two maps exhibit important differences: rates recorded in the vicinity of urban areas (i.e. low rates around Reno and Phoenix, or high rates around Eureka) are smoothed less by ATA kriging. Conversely,
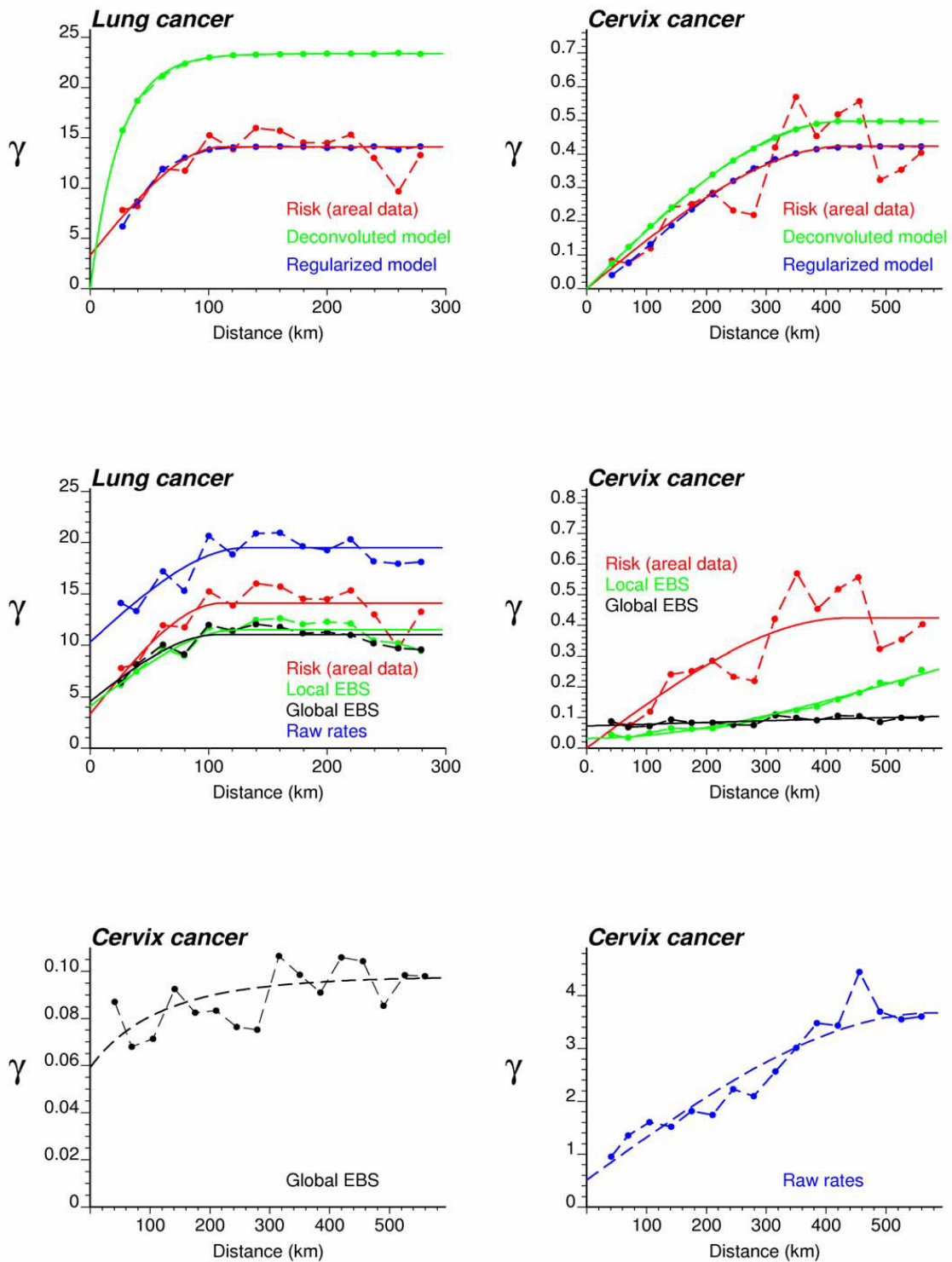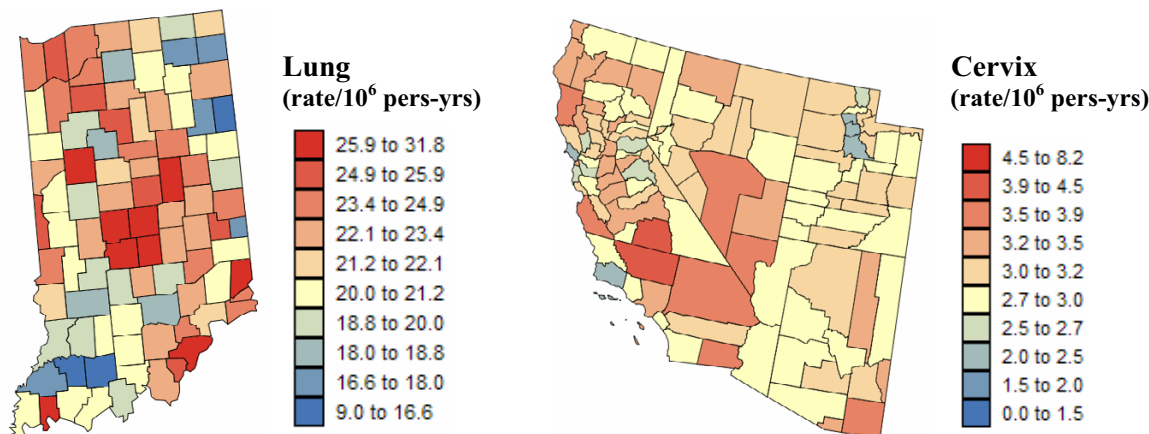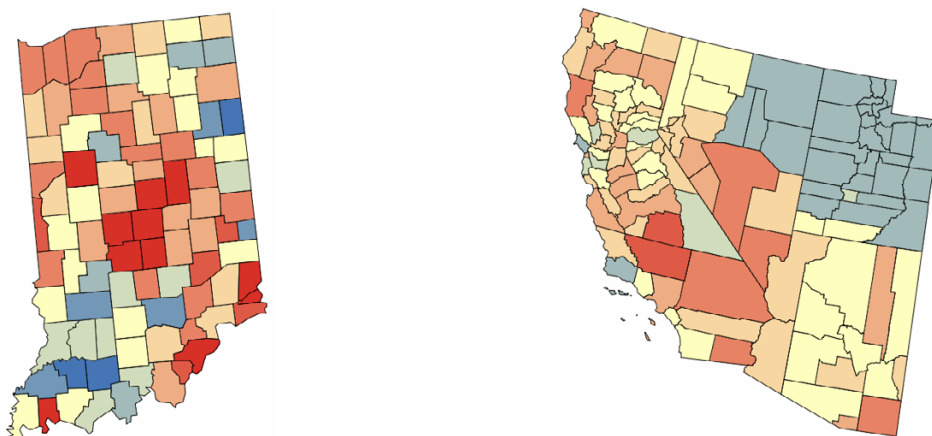
Figure 3
Semivariogram models used by the different geostatistical approaches for mapping lung and cervix cancer mortality. The top graphs show the experimental semivariogram of the risk estimated from county-level rate data (red curve), and the results of its deconvolution (green curve). The regularization of the point-support model yields a curve (blue line) that is very close to the experimental one. Bottom graphs show the semivariogram of raw rates, as well as the semivariograms of global and local empirical Bayes estimates.
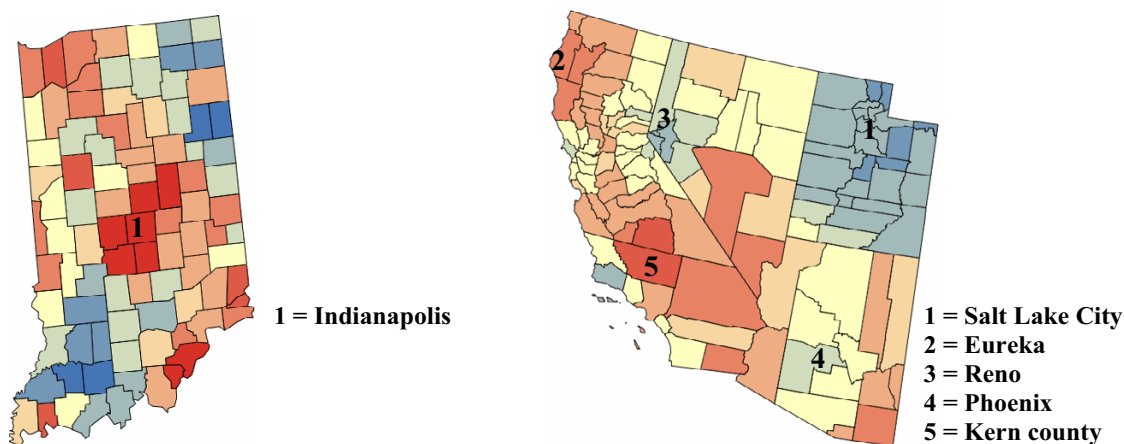
Figure 4
County maps of lung and cervix cancer mortality risks computed using alternative estimators. The fill color represents the risk estimated using the following approaches: global and local empirical Bayes smoothers (EBS), and area-to-area (ATA) Poisson kriging. The color legend applies to all the maps of the same region; the class boundaries correspond to the deciles of the histogram of original rates.

Table 1: Summary statistics for county-level and point-level estimates of lung and cervix cancer mortality.

Estimators	Lung cancer			Cervix cancer		
County-level estimates	Mean	Variance	Min-max	Mean	Variance	Min-Max
Observed rates	21.19	18.48	9.071–31.79	2.851	2.446	0.000–8.138
Global EBS	21.95	10.57	13.91–31.68	3.052	0.092	2.281–4.051
Local EBS	22.10	10.79	13.36–31.71	2.838	0.224	2.041–4.014
ATA Poisson kriging	21.58	10.96	13.22–31.56	2.851	0.264	1.921–4.081
Point estimates						
Kriged rates	21.15	4.970	15.46–26.85	2.972	1.003	0.470–6.702
Kriged global EBS	21.90	3.703	17.09–27.92	3.085	0.010	2.731–3.444
Kriged local EBS	22.02	4.520	16.25–28.24	2.883	0.154	2.061–3.519
ATP Poisson kriging	21.22	9.510	12.13–33.90	2.985	0.266	1.906–4.308
Aggregates of point estimates						
Kriged rates	21.22	4.832	16.14–26.29	2.849	0.641	0.955–5.084
Kriged global EBS	21.99	3.529	17.74–27.11	3.052	0.011	2.765–3.314
Kriged local EBS	22.12	4.390	16.94–27.43	2.842	0.157	2.064–3.406
ATP Poisson kriging	21.58	10.96	13.22–31.56	2.851	0.264	1.921–4.081

Mean, variance and range of cancer mortality rates per 100,000 person-years estimated at the county level or at the nodes of a 5 km grid using four alternative methods: point kriging of observed rates, point kriging of global and local empirical Bayesian smoothed (EBS) rates, area-to-area (ATA) and area-to-point (ATP) Poisson kriging. Statistics for the county-level aggregates of point estimates are listed at the bottom of the table.

more smoothing occurred in sparsely populated Eastern Nevada, leading to higher risk on the ATA kriged map versus local EBS. In summary, ATA kriging seems to weigh more heavily the population size in the filtering of the noisy rate maps.

Point-level estimates of mortality risk
To eliminate the visual bias associated with the interpretation of the choropleth maps of Figure 4, continuous risk surfaces were created for lung and cervix cancer mortality. The most straightforward mapping techniques amount at conducting a point kriging of the county-level raw or smoothed rates, which are simply assigned to the population-weighted centroids of those counties. In other words, one implicitly assumes that all the habitants of a county live at the same location and the measured rate thus refers to this specific location. Note that to ensure that both ATP and point kriging incorporate information on the spatial distribution of county population, the population-weighted centroids displayed in Figure 2 (middle maps) were used for semivariogram estimation and point kriging throughout the analysis. These centroids will be referred to as county centroids, as opposed to geographic centroids, hereafter. A prerequisite for kriging is a semivariogram model that is fitted to experimental values computed

from centroid-based rate data using estimator (25). Figure 3 (bottom graphs) shows that the largest semivariogram values are observed for raw rates (blue curve), which is expected because of the additional random variability caused by the small number problem. Since cervix cancer is less frequent than lung cancer, its mortality rates are more likely to be impacted by the small number problem and display higher levels of noise, leading to a sill that is one order of magnitude larger than for the other curves (Figure 3, right bottom graph). The semivariogram of empirical Bayes estimates has the smallest values since the smoothing reduces the variance of raw rates, in particular for cervix cancer. The use of local versus global empirical smoothers has little impact on the semivariogram for the frequent lung cancer while, as expected, the global smoothers lead to a much lower sill for the semivariogram of cervix cancer. For that cancer, the semivariogram of local EBS risk estimates (green curve) has a concave shape which reflects the presence of a trend that is apparent on the smoothed risk map of Figure 4 (i.e. low rate in Utah and high rates in Southern California). This semivariogram was modelled using a cubic model [6,31] with a range of 2,041 km, while the semivariogram model for local EBS estimates consists of two exponential models and a large nugget effect; see Figure 3 (left bottom graph).

Figures 5 and 6 show the risk values estimated at the nodes of a 5 km grid using the following methods: point kriging of raw rates, point kriging of global and local empirical Bayesian smoothed rates (EBS), and area-to-point Poisson kriging (ATP). In all four cases, the estimation was based on $K = 32$ closest county-level rates (local search window). To enforce the coherence constraint, within a given county each grid node was estimated using the same 32 county-level rates; that is the proximity is measured in terms of distance between the centroids of both the prediction and the data counties instead of the distance between the grid node in the prediction county and the centroid of the data county. For lung cancer, the maps created by point kriging of raw and EBS rates show only minor differences; higher EBS risk estimates are found close to the Northeast and Northwest borders where low rates in sparsely populated counties were smoothed by the empirical Bayes method. The ATP kriged map displays much more details, with the presence of clearly delineated areas of lower and higher rates. These visual impressions are confirmed by the statistics listed in Table 1: the variance of the set of ATP kriging estimates (9.510) is twice the variance calculated for the other methods (3.703–4.970). Similar conclusions can be drawn from the comparison of the ATP and EBS kriging maps for Region 2: the variance of ATP kriging estimates (0.266) is twice the variance obtained for the kriging of local EBS (0.154) and more than one order of magnitude larger than for the global empirical Bayes smoother (0.010). The isopleth map of GBS risk estimates is very smooth because of the large nugget effect of the corresponding semivariogram model (recall Figure 3, left bottom graph). Conversely, the kriged map of raw rates is the most variable since it is based on the interpolation of unstable cervix cancer mortality rates, and one can distinguish very clearly the location of several county centroids.

Although the interpretation of spatial patterns is beyond the scope of this methodology paper, comparison of ATP and local EBS kriging maps for Region 2 offers interesting insights about the spatial distribution of cervix cancer mortality risk:

• The impact of low or high rates recorded in the vicinity of urban areas (i.e. low rates around Reno and Phoenix, or high rates around Eureka) is much more apparent on the ATP kriged map. This short-scale variability in the risk map results from the greater weight assigned to population density and areal data through the coherence constraint in area-to-point kriging.

• Lower cervix cancer mortality is clearly confined to Utah on the ATP kriging map, while these low rates expand to Eastern Nevada on the maps of kriged local EBS.
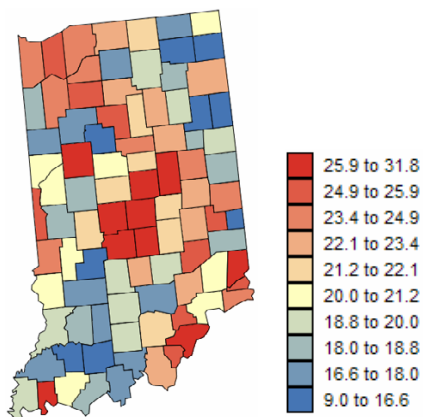
• ATP kriging highlights a cluster of high mortality risk in Southern California. A similar cluster was identified with Kern County on the choropleth map of ATA kriged risk in Figure 4. The isopleth risk map shows that high risks are not confined to this sole county but potentially spread over four counties. This information, which is important for designing prevention strategies, is blurred by the use of county-level estimates.

Figures 7 and 8 show the maps of the kriging variance associated with the risk estimates of Figures 5 and 6. Differences among interpolation methods are much more pronounced in terms of prediction variance than estimated risk. The point kriging variance is mainly influenced by the location of county centroids (i.e. data geometry): prediction variances are small in the vicinity of county centroids and increase in extrapolation situation, e.g. close to the state borders for the maps of Region 1 (Figure 7). The location of county centroids is particularly obvious on the global EBS map in Region 2 because of the large nugget effect of the corresponding semivariogram. Furthermore, the variance map for local EBS is very smooth as a result of the highly continuous behaviour of the cubic semivariogram model at the origin; recall Figure 3 (right middle graph, green curve). The pattern of these point kriging variance maps essentially reflects the presence of two clusters of small counties (i.e. nearby county centroids) in Northern California and in Utah. The point kriging variance indirectly accounts for the spatial distribution of population within each county through the use of population-weighted centroids, yet differences in population sizes among counties is ignored. This information is incorporated directly into the computation of the ATP kriging variance leading to increased uncertainty in sparsely populated areas. This effect is particularly apparent in Region 2; compare the bottom right map in Figure 8 with the top population map. One can easily associate the location of pockets of low variance with urban centers, such as Los Angeles, San Francisco, Salt Lake City, Las Vegas or Tucson. In both regions, the variance of EBS point risk estimates is the smallest since the smoothing leads to semivariograms with the lowest sills; recall Figure 3.
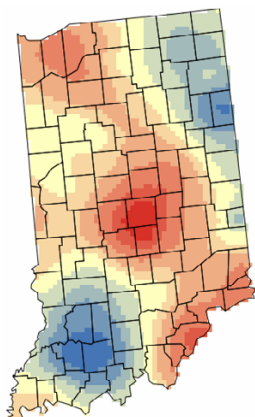
Coherence of estimation techniques
The isopleth risk maps displayed in Figures 5 and 6 can be viewed as the product of the disaggregation of the choropleth maps of Figure 4. One should thus expect that the aggregation of point kriging risk estimates within each county returns the areal data for that county. This "coherence constraint" is not implicit to the point kriging approach, and Table 1 (bottom rows) confirms that the distributions of aggregated point risk estimates and of the county-level rates used in point kriging do not share the same statistics. Because of the smoothing effect of kriging,
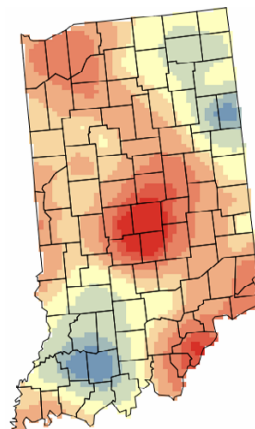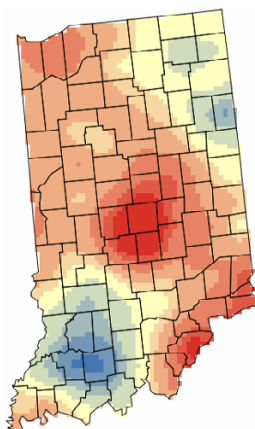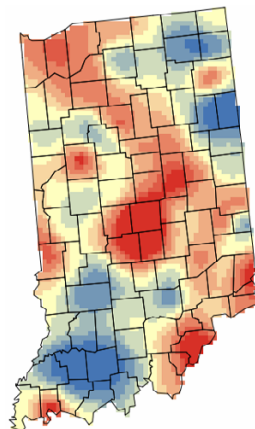
Figure 5
Map of lung cancer mortality rates in Region 1, and the risk estimated by the different forms of kriging. The fill color represents the age-adjusted mortality rate per 100,000 person-years recorded over the period 1970–1994 (top graph) or the risk estimated using the following approaches: point kriging of raw rates, point kriging of global and local empirical Bayes smoothers (EBS), and area-to-point (ATP) Poisson kriging. The color legend applies to all the maps; the class boundaries correspond to the deciles of the histogram of original rates.
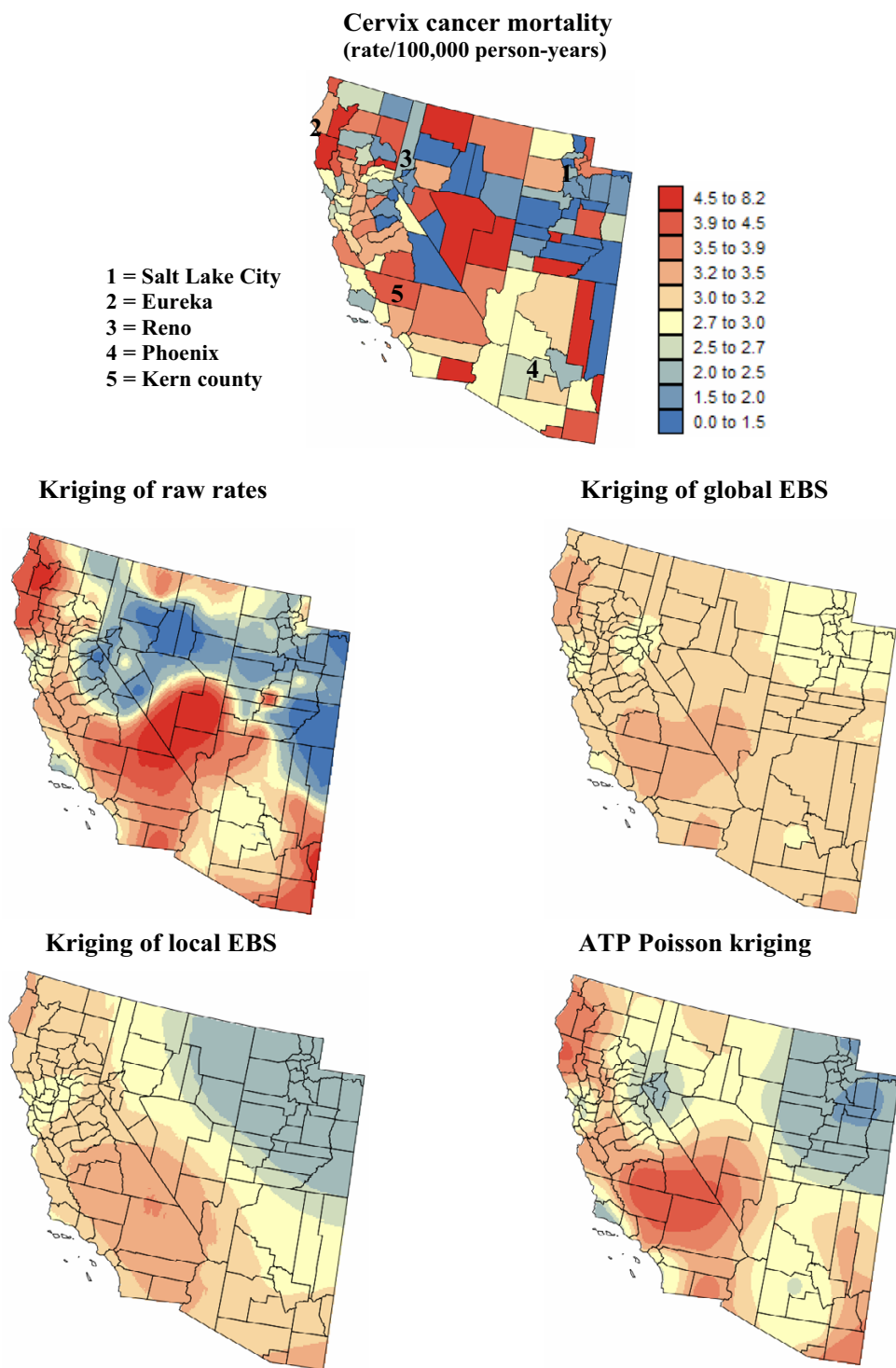
Figure 6
Map of cervix cancer mortality rates in Region 2, and the risk estimated by the different forms of kriging. The fill color represents the age-adjusted mortality rate per 100,000 person-years recorded over the period 1970–1994 (top graph) or the risk estimated using the following approaches: point kriging of raw rates, point kriging of global and local empirical Bayes smoothers (EBS), and area-to-point (ATP) Poisson kriging. The color legend applies to all the maps; the class boundaries correspond to the deciles of the histogram of original rates.
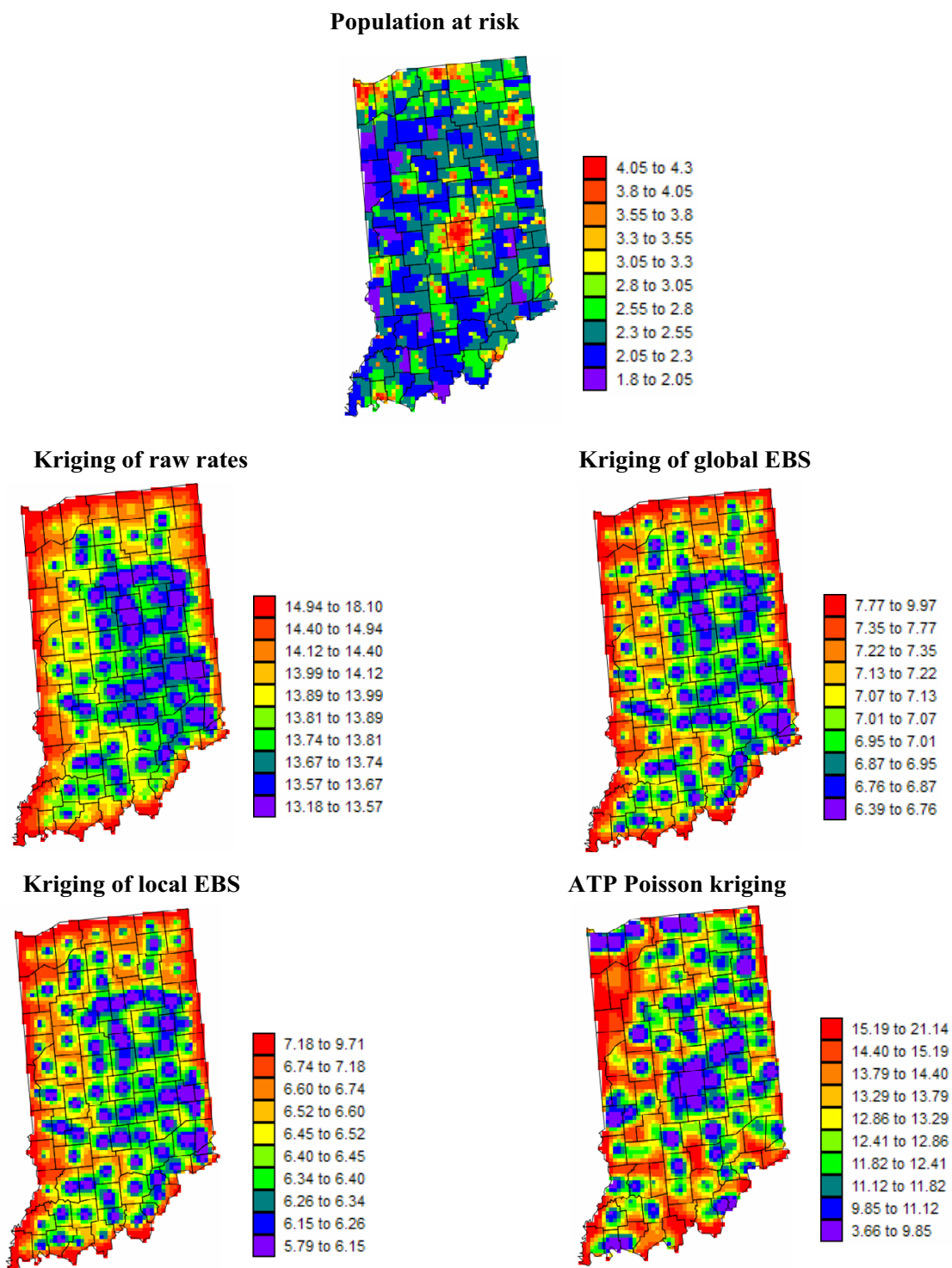
Figure 7
White female population map for Region 1, and the prediction variance for the different forms of kriging. The fill color represents the population at risk (lognormal scale) or the kriging variance associated with the risk maps of Figure 5. The following estimation techniques were used: point kriging of raw rates, point kriging of global and local empirical Bayes smoothers (EBS), and area-to-point (ATP) Poisson kriging. The units for the kriging variance maps are (age-adjusted mortality rate per 100,000 person-years)^2, and the class boundaries correspond to the deciles of the histogram of variance.
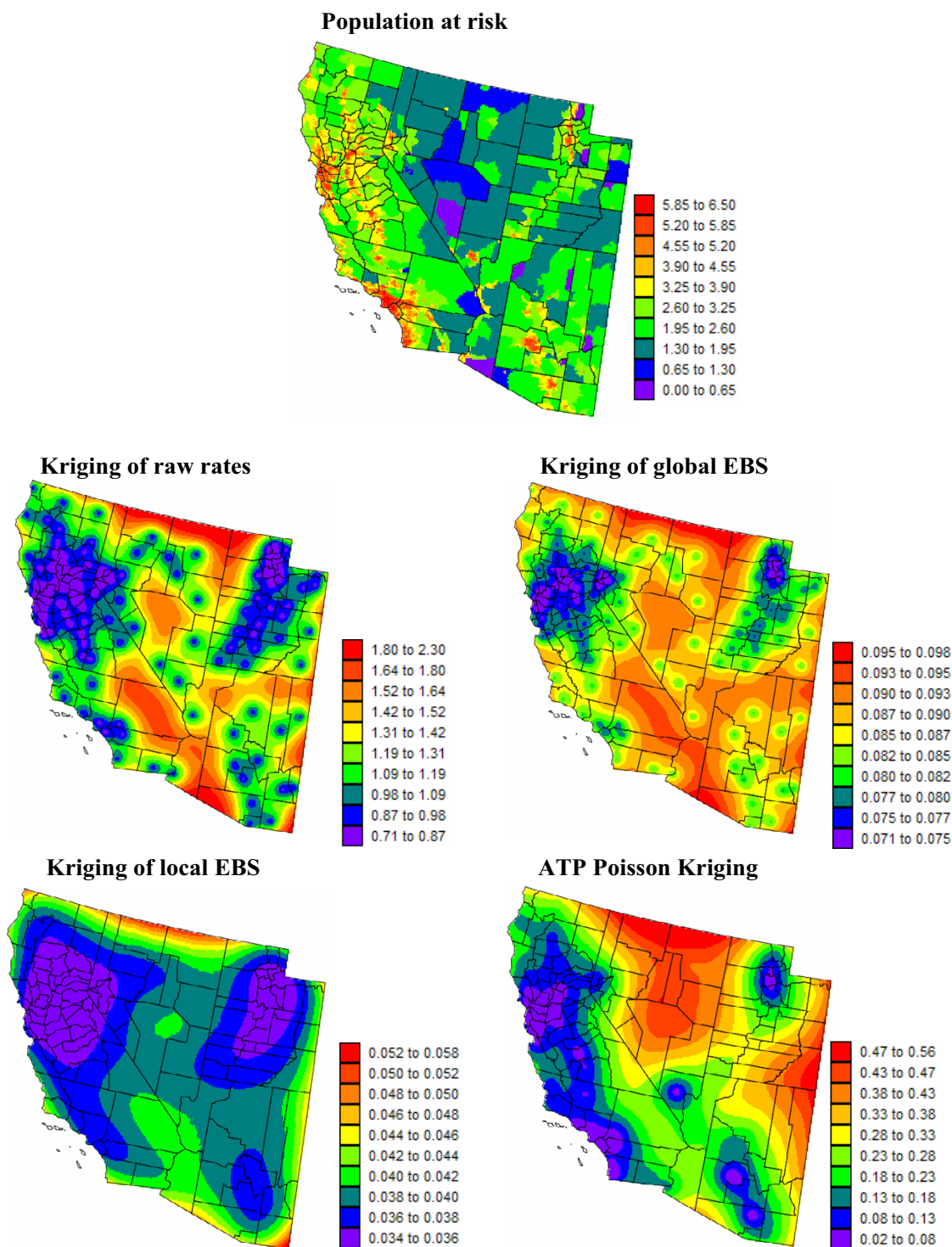
Figure 8
White female population map for Region 2, and the prediction variance for the different forms of kriging. The fill color represents the population at risk (lognormal scale) or the kriging variance associated with the risk maps of Figure 6. The following estimation techniques were used: point kriging of raw rates, point kriging of global and local empirical Bayes smoothers (EBS), and area-to-point (ATP) Poisson kriging. The units for the kriging variance maps are (age-adjusted mortality rate per 100,000 person-years)^2, and the class boundaries correspond to the deciles of the histogram of variance.

the variance of aggregated estimates is 50 to 80% smaller than the variance of areal data. This is not the case for ATP kriging where the coherence constraint (Equation 15) ensures that the population-weighted variance, as well as the population-weighted mean, of ATP kriging estimates equals the variance and mean of ATA kriging estimates; see Table 1. This coherence constraint explains the greater level of details noticed in the ATP kriging map versus the risk maps created by the simple point kriging techniques.

Point Poisson kriging versus ATA Poisson kriging
In earlier papers on the use of Poisson kriging for filtering noise in cancer mortality maps [15], a point kriging approach was implemented whereby only county geographic centroids were considered for semivariogram estimation and kriging. To investigate the impact of these simplifying assumptions on the computation of the kriging estimate and variance the results of ATA and point Poisson kriging are compared in Figure 9. Estimation of the mortality risk is fairly robust with respect to the use of centroid geography, in particular for Region 1 where all counties have similar size and shape. Discrepancies between both approaches are much more pronounced for the kriging variance: ignoring the spatial support of county-level data leads almost systematically to larger kriging variance; see Figure 9 (bottom graphs). This result is consistent with the "change-of-support" theory that predicts smaller error variances for block versus point estimation. The magnitude of the overestimation by point kriging is the largest for Region 2 that includes very large counties and has a wide range of county sizes and shapes.

Simulation studies
Figures 3 through 8 illustrated the major differences between alternative approaches for creating isopleth risk maps from aggregated rates. An objective assessment of the prediction performances of the various techniques requires, however, the availability of the underlying risk maps, which are unknown in practice. Simulation provides a way to generate multiple realizations of the spatial distribution of cancer mortality rates that can be analyzed using alternative approaches. Predicted risks can then be compared to the risk maps used in the simulation.

For both lung and cervix cancers, $L = 100$ maps of county-level mortality rates $\{z^{(l)}(v_\alpha), \alpha = 1,..., N; l = 1,..., L\}$ were simulated using the following procedure:

1. Continuous maps of mortality risk values, $\{r^{(l)}(u_s), s = 1,..., S\}$, are first simulated using non-conditional sequential Gaussian simulation (see [16,9], p. 380 for a description of the algorithm). The simulation grid, which has a 5 km spacing, consists of $S = 3,751$ and $S = 48,474$ nodes for Regions 1 and 2, respectively. Five different risk maps (i.e. $L = 5$) were generated for each cancer, and two realizations

for each region are displayed at the top of Figures 10 and 11. Each realization reproduces the histogram of raw cancer mortality rates, while the deconvoluted semivariogram models of Figure 3 were used in the simulation algorithm: exponential model with a range of 75 km for Region 1 and a spherical model with range of 425 km for Region 2. Both models have a zero nugget effect and a unit sill.

2. Each simulated risk map was combined with the population maps of Figure 2 to compute a number of deaths for all grid nodes. Both simulated death counts and population data were then aggregated within each county, and their ratio defines the simulated county-level mortality risk:

$$r^{(l)}(v_\alpha) = \frac{1}{n(v_\alpha)} \sum_{s=1}^{P_\alpha} n(u_s) r^{(l)}(u_s) \quad \text{with} \quad n(v_\alpha) = \sum_{s=1}^{P_\alpha} n(u_s) \qquad (\text{Equation } 30)$$

The maps of aggregated risk values, $\{r^{(l)}(v_\alpha), \alpha = 1,..., N\}$, are displayed in the middle of Figures 10 and 11.

3. For each cancer and each of the five risk maps, 20 realizations of the number of deaths recorded over each county v_α were generated by random drawing of a Poisson distribution whose mean parameter is $r^{(l)}(v_\alpha) \times n(v_\alpha)$. The division of the simulated death counts by the county population leads to $5 \times 20 = 100$ sets of simulated mortality rates for each cancer; see bottom of Figures 10 and 11 for the first realization generated for the risk maps displayed in the same figures. The noise caused by the small number problem is particularly apparent for cervix cancer. For example, a couple of counties in the North central part of realization #1 display high mortality rates, while the underlying risk value is low.

Comparison of prediction performances
Each map of simulated rates $\{z^{(l)}(v_\alpha), \alpha = 1,..., N\}$ underwent a (geo)statistical analysis in order to estimate point risk values using four alternate approaches: point kriging of raw rates, point kriging of global and local empirical Bayesian smoothed rates (EBS), and area-to-point Poisson kriging (ATP). In all four cases, the estimation at each grid node was based on $K = 32$ closest areal data (local search window). For each realization, the semivariogram needed for each type of kriging was estimated before being automatically modelled and deconvoluted in the case of ATP kriging.

Figure 12 shows the different semivariogram models for the two simulated maps of lung and cervix cancer mortality displayed at the bottom of Figures 10 and 11. The reference point (black curve) and areal (yellow curve) support models are the semivariograms fitted directly to the original simulated risk maps before and after aggrega-
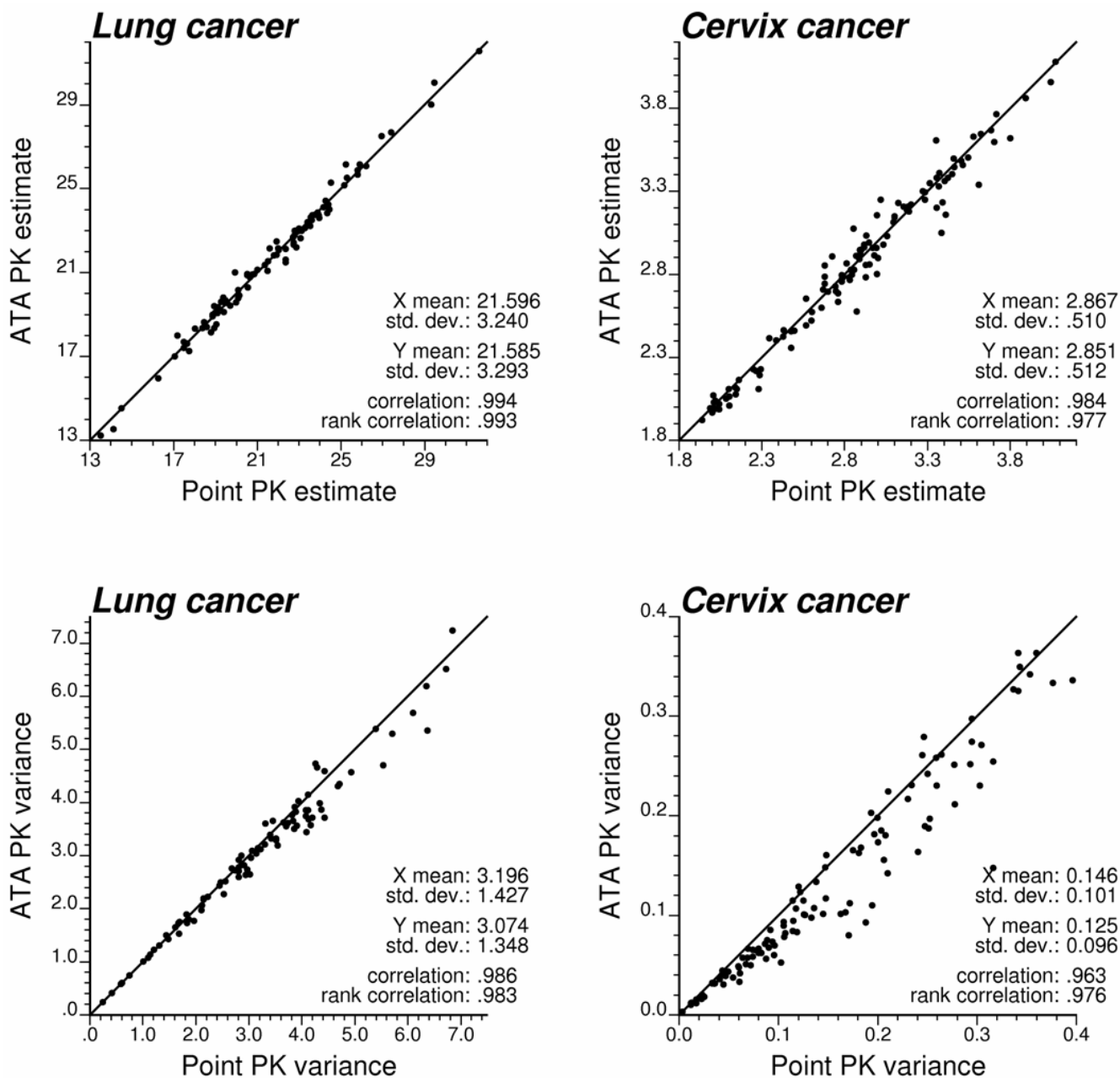
Figure 9
Impact of ignoring the spatial support and population distribution on Poisson kriging results. Top scatterplots illustrate the similarity of the cancer mortality risk estimated using punctual and area-to-area (ATA) Poisson kriging. Ignoring the shape and size of counties (i.e. collapsing county rates to their population-weighted centroids) leads to an overestimation of the kriging variance, in particular for the set of vastly different counties in Region 2.

tion to the county level; see Figure 12 (left column). Models resulting from the deconvolution procedure are plotted as green curves. Comparison of the green and black solid curves indicates that the deconvolution yields point-support models that are reasonably close to the underlying ones in terms of range values. Each deconvo-

luted model was regularized according to Equation (18), and the resulting model (blue curve) agrees fairly well with the risk semivariogram model fitted to areal data (red curve), which demonstrates the convergence of the iterative deconvolution procedure. Note also how this risk semivariogram model is close to the reference areal sup-
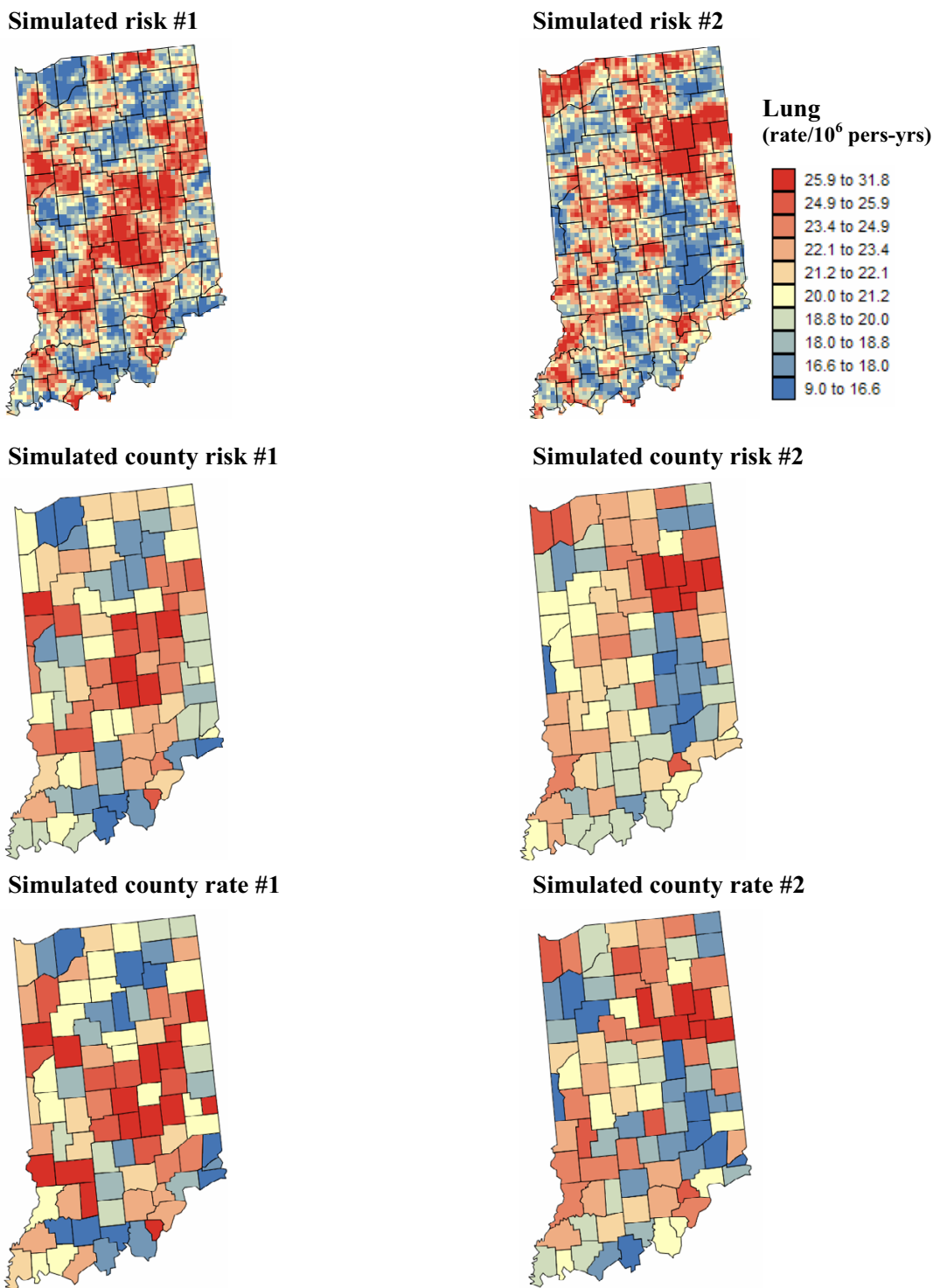
Figure 10
County maps of lung cancer mortality rates simulated under two scenarios for the underlying continuous risk map. The number of cases for each county was simulated by random sampling of a Poisson distribution that is defined by the white female population map of Figure 1 and the county-level aggregation of a continuous risk map generated using sequential Gaussian simulation. The units are age-adjusted mortality rates per 100,000 person-years. The color legend applies to all the maps; the class boundaries correspond to the deciles of the histogram of original rates.

Figure 11
County maps of cervix cancer mortality rates simulated under two scenarios for the underlying continuous risk map. The number of cases for each county was simulated by random sampling of a Poisson distribution that is defined by the white female population map of Figure 1 and the county-level aggregation of a continuous risk map generated using sequential Gaussian simulation. The units are age-adjusted mortality rates per 100,000 person-years. The color legend applies to all the maps; the class boundaries correspond to the deciles of the histogram of original rates.
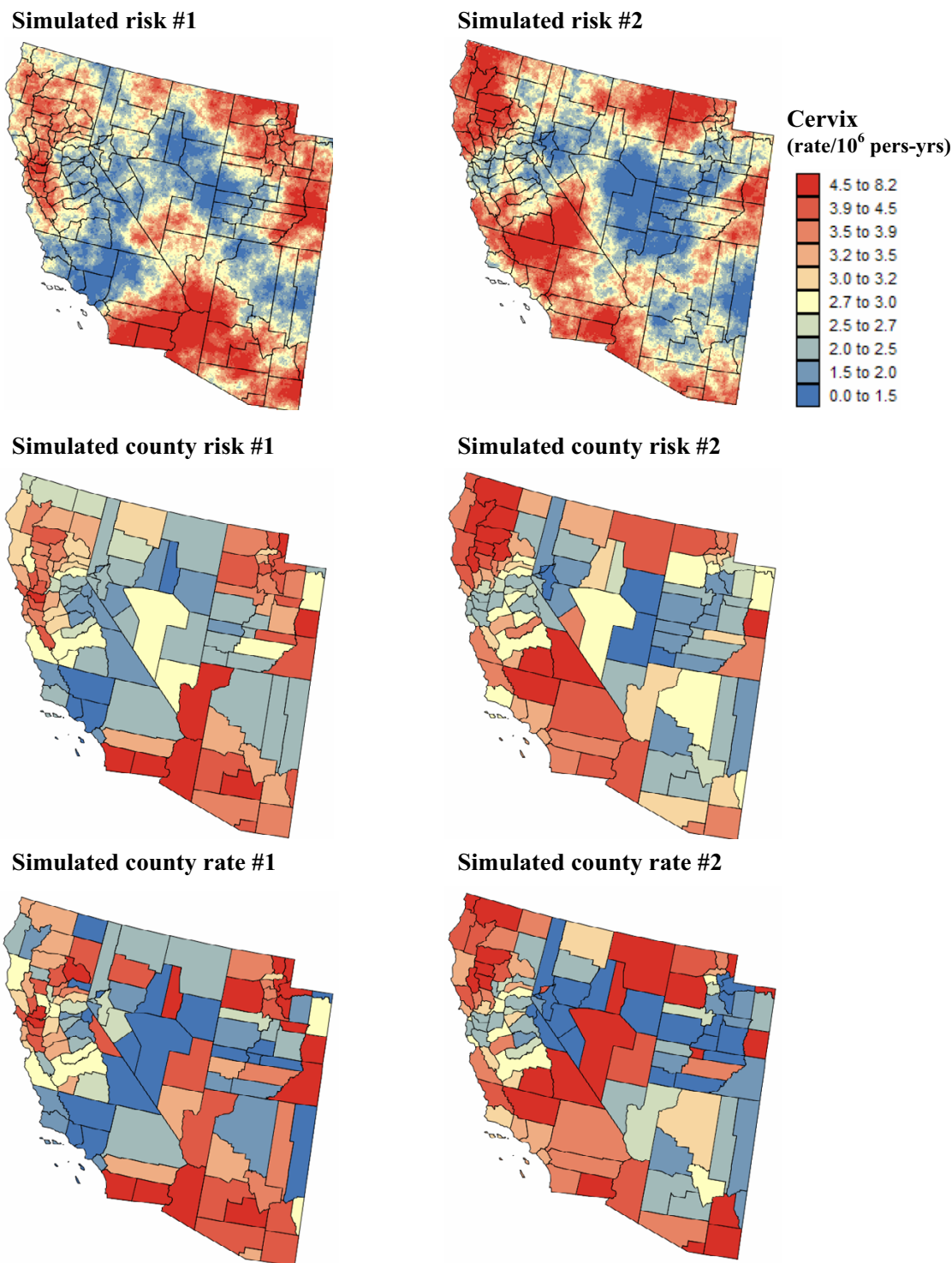
port model (yellow curve). Because of the good agreement between theoretically regularized and the risk semivariogram models inferred from the county-level rates, discrepancies at the point-support level are essentially caused by the use of the regularization formula (Equation 18). An important factor that influences the deconvolution results is the behavior at the origin of the regularized and point-support semivariogram models; for example the fitting of a nugget effect or the use of a spherical (linear behavior) versus cubic (parabolic behavior) model. Unfortunately, in absence of point data this part of the semivariogram model can not be validated. In this paper, no nugget effect was systematically fitted to the point-support model to account for the characteristics of the model used in the simulation.

The semivariogram models used by the point kriging approaches are plotted in the right column of Figure 12. In particular for cervix cancer, the noise caused by the small number problem leads to semivariograms for raw rates that have a high sill and a large relative nugget effect (blue curves). Rate smoothing using global or local empirical Bayes approaches substantially reduces the sample variability, resulting in smaller sills and nuggets effects, as well as parabolic behaviour for most of the semivariograms at the origin (green and black curves). The risk semivariogram estimator (Equation 5), which accounts for population size in the estimation, yields models with intermediate sills (red curves).

Figures 13 and 14 show, for the simulated rate maps #1 of Figures 10 and 11, the risk values estimated at the nodes of a 5 km grid using the four alternate methods. The analysis of simulated maps confirms the conclusions drawn from the analysis of real cancer mortality rates in Figures 5 and 6. The location of county centroids is the most apparent on the kriged map of raw rates which is also the most variable because it is based on the interpolation of unstable mortality rates, in particular for cervix cancer. Kriging of global and local EBS estimates leads to very smooth maps free of the influence of unreliable rates. Poisson kriging, through its coherence constraint, generates maps of risk estimates that are locally more detailed (i.e. less smoothing) and reveal features of the true risk maps that were blurred by empirical Bayes smoothing, such as the low risk area extending up to the eastern border of Region 2.

For all 100 realizations of each type of cancer, predicted risks $\{ r_P^{(l)}(\mathbf{u}_s), s = 1,..., S\}$ and the corresponding prediction variance $\{[\sigma_P^{(l)}(\mathbf{u}_s)]^2, s = 1,..., S\}$ were compared to the underlying risk map $\{r(\mathbf{u}_s), s = 1,..., S\}$. As in a previous study on the performances of alternative smoothing techniques [15], various criteria were computed and averaged over all realizations to attenuate the impact of statistical fluctuations.

Bias and accuracy of prediction
The first two criteria are the mean error (ME) and mean absolute error (MAE) of prediction computed as:

$$ME^{(l)} = \frac{1}{W} \sum_{s=1}^{S} \omega_s \left[r_P^{(l)}(\mathbf{u}_s) - r(\mathbf{u}_s) \right] \quad \text{with} \quad W = \sum_{s=1}^{S} \omega_s \qquad (\text{Equation } 31)$$

$$MAE^{(l)} = \frac{1}{W} \sum_{s=1}^{S} \omega_s \left| r_P^{(l)}(\mathbf{u}_s) - r(\mathbf{u}_s) \right| \quad \text{with} \quad W = \sum_{s=1}^{S} \omega_s \qquad (\text{Equation } 32)$$

The prediction error at each grid node $\mathbf{u}_s$ is weighted according to the population size at that location, $\omega_s = n(\mathbf{u}_s)$, in order to penalize more the errors that affect a larger population. Table 2 (top rows) indicates that, on average over 100 realizations, all prediction methods are relatively unbiased. However, in particular for cervix, the smallest bias is the most frequently (i.e. 51% of realizations) observed for ATP kriging estimates. Methods differ much more in terms of the mean absolute error of prediction; see Table 2 (bottom rows). On average, ATP kriging reduces the absolute error of simple prediction methods by 20% for cervix and 5–10% for lung. The benefit of Poisson kriging over kriging of raw or empirical Bayesian smoothed rates is almost systematic since it leads to smaller MAE values 99% of the time for cervix and 89% for lung cancer. Although the deconvoluted semivariogram models depart somewhat from the true point-support models (recall Figure 12, left column), the use of the true model $\gamma_R(\mathbf{h})$ (black curve in Figure 12, left column) in ATP kriging does not cause a substantial decline in prediction errors; compare the last two rows of Table 2 for each performance criterion.

Variance of the prediction errors and smoothing effect
In addition to a risk estimate, all interpolation methods provide a measure of the uncertainty attached to this estimate in the form of the kriging variance (Equations 12 and 23). For the point kriging approaches, the variance $[\sigma_P^{(l)}(\mathbf{u}_s)]^2$ depends on the proximity of county centroids (i.e. center of mass of population within the county) to the grid node $\mathbf{u}_s$, as well as the form of the semivariogram model (i.e. nugget effect, range of autocorrelation). The ATP kriging variance directly accounts for the population size at the county and grid levels. For each interpolation method, the kriging variance was averaged over all S grid nodes, resulting in the following statistic:
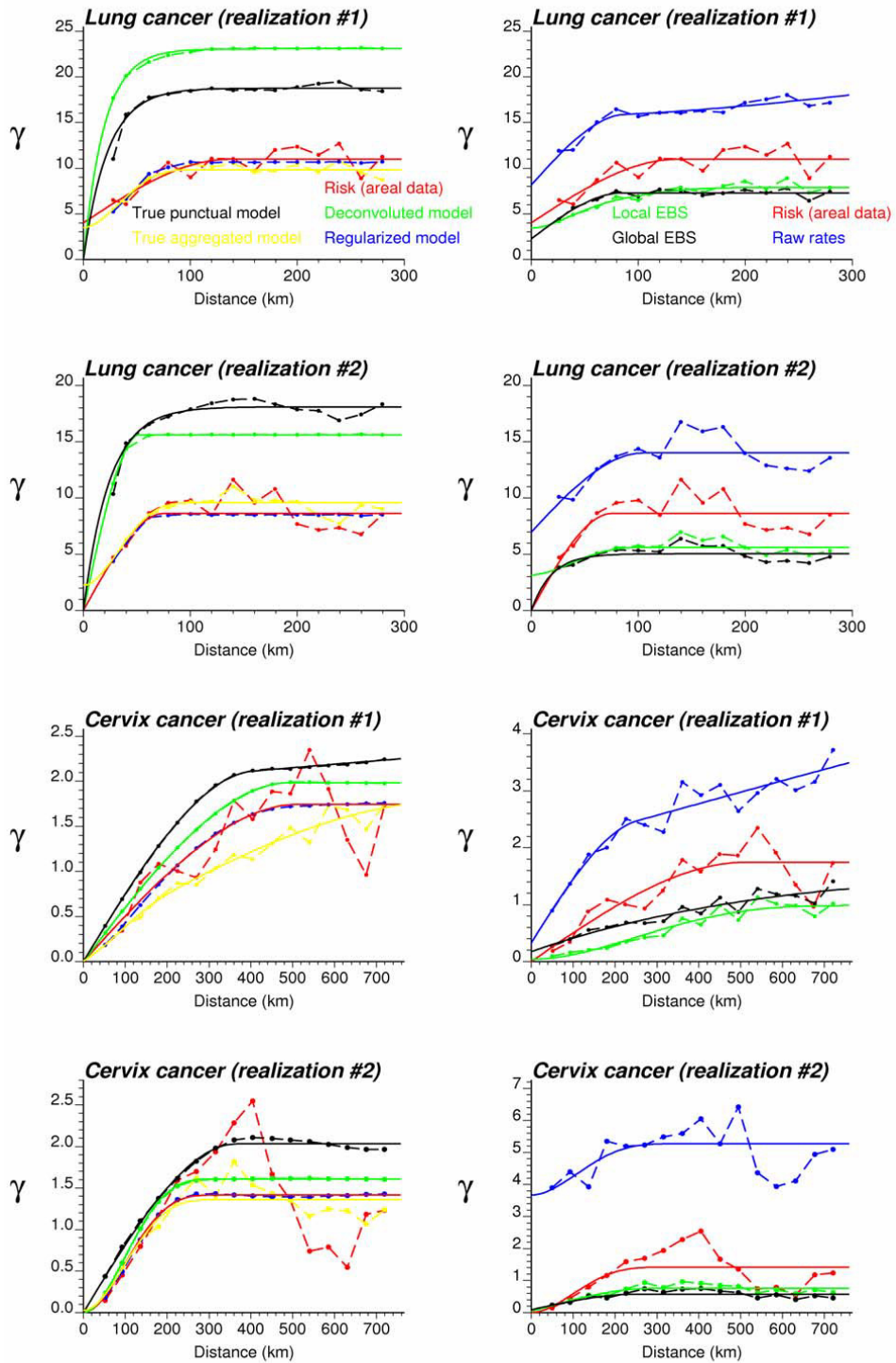
Figure 12
Semivariogram models inferred from the simulated rate maps of Figures 10 and 11. The left column shows the experimental semivariogram of the risk estimated from county-level rate data (red curve), and the results of its deconvolution (green curve) which is reasonably close to the "true" point-support model (black curve). The regularization of the point-support model yields a curve (blue line) that is very close to the experimental one. Yellow curves represent the semivariograms computed from the underlying county risk maps. Right column shows the semivariogram of raw rates, as well as the semivariograms of global and local empirical Bayesian smoothed (EBS) rates.
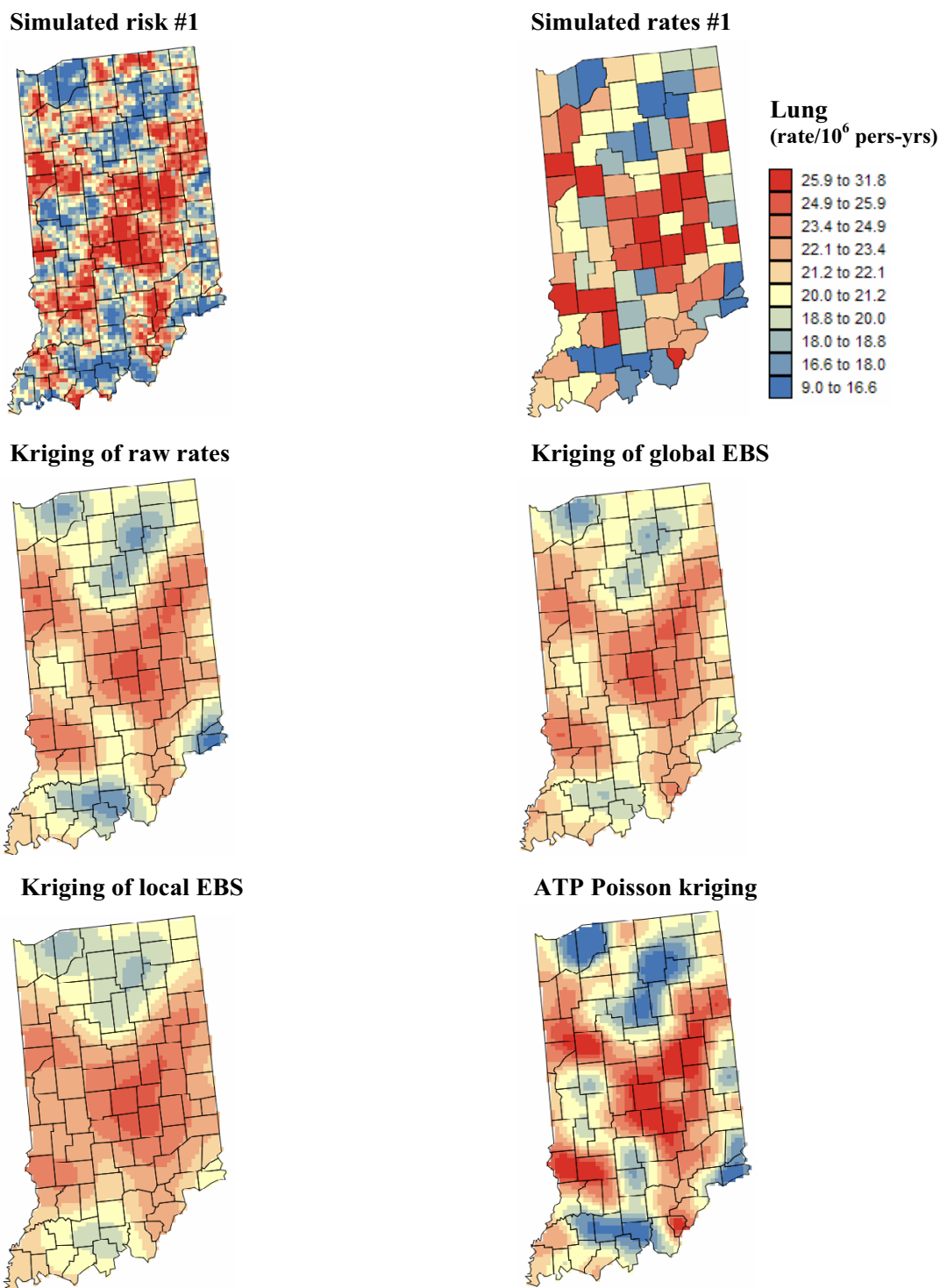
Figure 13
Simulated lung cancer risk map and the results of the different forms of area-to-point kriging. The fill color represents mortality risk per 100,000 person-years simulated by sequential Gaussian simulation and the results of the estimation using the following approaches: point kriging of raw rates, point kriging of global and local empirical Bayesian smoothed (EBS) rates, and area-to-point (ATP) Poisson kriging. The color legend applies to all the maps; the class boundaries correspond to the deciles of the histogram of original rates.

Figure 14
Simulated cervix cancer risk map and the results of the different forms of area-to-point kriging. The fill color represents mortality risk per 100,000 person-years simulated by sequential Gaussian simulation and the results of the estimation using the following approaches: point kriging of raw rates, point kriging of global and local empirical Bayesian smoothed (EBS) rates, and area-to-point (ATP) Poisson kriging. The color legend applies to all the maps; the class boundaries correspond to the deciles of the histogram of original rates.
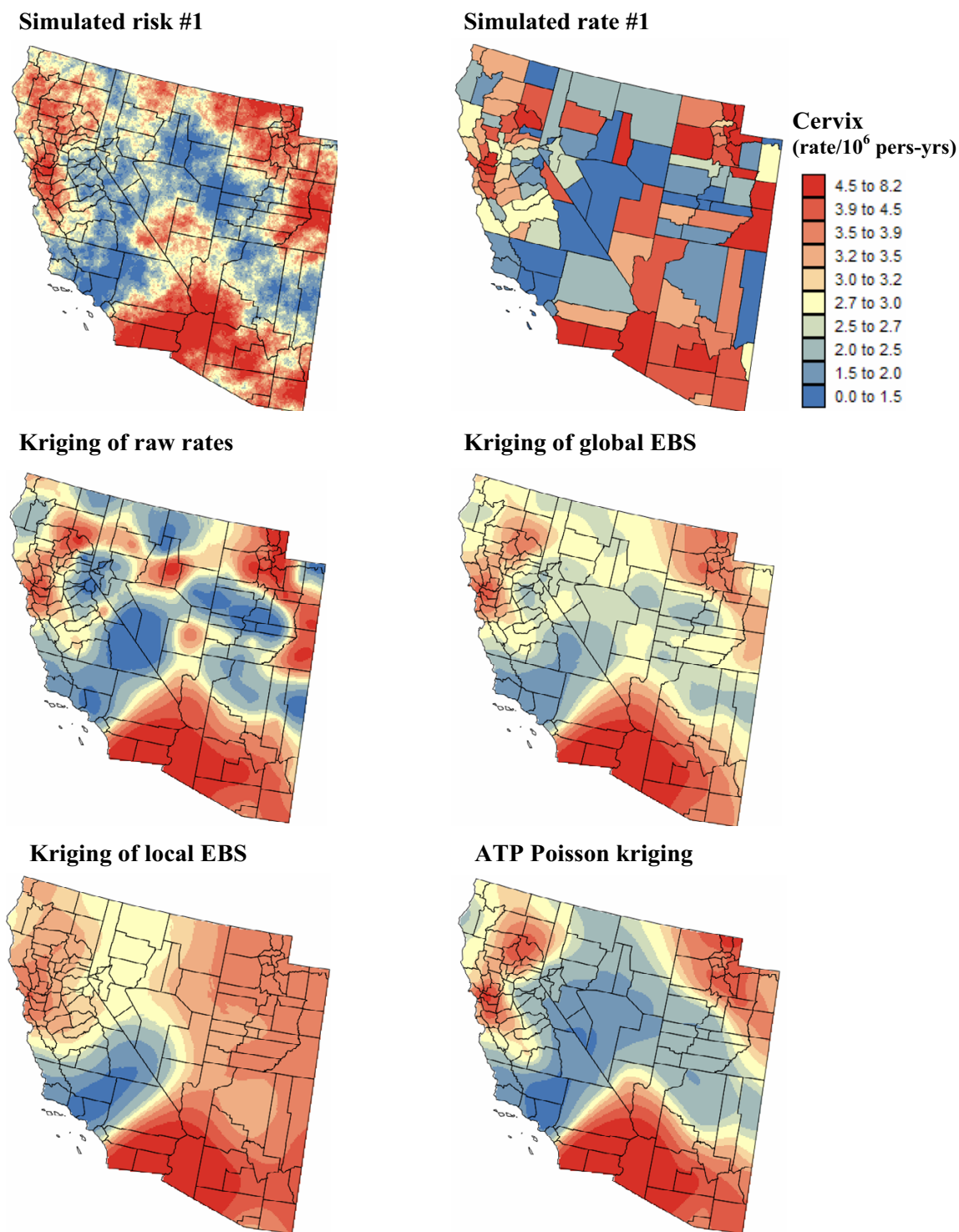
Table 2: Performance comparison of alternative kriging estimators: mean errors and mean absolute errors of prediction.

Estimators	Lung cancer		Cervix cancer	
MEAN ERROR	Average	% best result	Average	% best result
Point kriging of raw rates	-0.013	41	0.060	11
Point kriging of global EBS	**0.008**	11	0.040	15
Point kriging of local EBS	0.042	8	0.044	23
ATP Poisson kriging	-0.036	40	**-0.001**	51
ATP Poisson kriging (true $\gamma_R(\mathbf{h})$)	-0.032		-0.001	
MEAN ABSOLUTE ERROR				
Point kriging of raw rates	2.647	8	0.406	0
Point kriging of global EBS	2.694	3	0.406	1
Point kriging of local EBS	2.776	0	0.418	0
ATP Poisson kriging	**2.471**	89	**0.317**	99
ATP Poisson kriging (true $\gamma_R(\mathbf{h})$)	2.452		0.313	

Results obtained on average over 100 realizations generated for Regions 1 and 2. Poisson kriging was conducted with the semivariogram model derived through deconvolution or inferred directly from the simulated grid risk values (true point-support model $\gamma_R(\mathbf{h})$). Bold numbers refer to best performances outside the ideal case where the true semivariogram of risk is known. The second column gives the percentage of realizations where the particular method (except ATP kriging with true $\gamma_R(\mathbf{h})$) yields the smallest prediction error.

$$\mathrm{VPE}^{(l)} = \frac{1}{S}\sum_{s=1}^{S}\left[\sigma_P^{(l)}(\mathbf{u}_s)\right]^2 \qquad (\text{Equation 33})$$

For both cancers, the smallest kriging variance is obtained for point kriging of smoothed (EBS) rates (see Table 3), which is expected because of the lower sill of the corresponding semivariogram models; recall examples of Figure 12 (right column). Conversely, the noise caused by the small number problem leads to higher sills for the semivariograms of raw rates, hence larger kriging variances, in particular for cervix cancer.

The semivariogram model not only influences the magnitude of the kriging variance at each grid node $\mathbf{u}_s$, but it also controls the smoothness of the maps of risk estimates. The variance of the risk estimates was computed for each realization and the average over 100 realizations is listed in Table 3. The variance of the reference risk values, averaged over the five scenarios for the risk map, is 18.137 and 1.828 for lung and cervix cancers, respectively. In agreement with the results displayed for one realization in Figures 13 and 14, point kriging of empirical Bayes estimates generates the largest smoothing effect: the variance represents 15–25% of the reference risk variance. The smoothing effect of ATP kriging is much smaller and similar for both cancers: the ratio of variances is 47–50%. Results for point kriging of raw rates are strongly affected by the reliability of the original data: the smoothing is much smaller when interpolating the unreliable cervix mortality rates (ratio = 94%) than for lung cancer (ratio = 31%).

Table 3: Performance comparison of alternative kriging estimators: kriging variance and smoothing effect.

Estimators	Lung cancer		Cervix cancer	
	Kriging variance	Dispersion variance	Kriging variance	Dispersion variance
Point kriging of raw rates	9.691	5.536	2.724	1.726
Point kriging of global EBS	3.423	2.934	0.314	0.281
Point kriging of local EBS	3.307	2.902	0.237	0.448
ATP Poisson kriging	9.378	8.401	0.645	0.926
ATP Poisson kriging (true $\gamma_R(\mathbf{h})$)	9.470	8.654	0.849	1.089

Results obtained on average over 100 realizations generated for Regions 1 and 2. The dispersion variance refers to the variance of risk estimates. Poisson kriging was conducted with the semivariogram model derived through deconvolution or inferred directly from the simulated grid risk values (true point-support model $\gamma_R(\mathbf{h})$).

Quality of the uncertainty model

According to the results of Table 3, point kriging of EBS rates yields the most accurate prediction since the corresponding kriging variance is the smallest; a conclusion that conflicts with results of Table 2. The ability of the prediction variance to capture the actual magnitude of the prediction error was quantified using the following mean square standardized residual (MSSR):

$$MSSR^{(l)} = \frac{1}{S} \sum_{s=1}^{S} \left[\frac{r_P^{(l)}(\mathbf{u}_s) - r(\mathbf{u}_s)}{\sigma_P^{(l)}(\mathbf{u}_s)} \right]^2 \qquad (\text{Equation 34})$$

If the actual estimation error is equal, on average, to the error predicted by the model, the MSSR statistic should be about one [[31], p. 91]. The MSSR statistic was averaged across the *L* realizations according to the following formula:

$$MSSR = \frac{1}{L} \sum_{l=1}^{L} \left[\delta_l MSSR^{(l)} + \frac{(1 - \delta_l)}{MSSR^{(l)}} \right] \qquad (\text{Equation 35})$$

with $\delta_l = 1$ if MSSR$^{(l)} > 1$, and zero otherwise. This averaging allows one to penalize equally the over- and underestimation of the prediction errors by the kriging variance. Table 4 (top rows) indicates that the EBS kriging variance fails to inform on the actual magnitude of the prediction errors. Because $\delta_l = 1$ for all realizations, the MSSR statistic means that, on average over all realizations, the actual squared prediction error is 4 to 7 times larger than what is predicted using the kriging variance. This over-optimistic assessment of the performance of EBS kriging results from the smoothing of rates in the empirical Bayes approach. For lung cancer, point kriging of raw rates yields the best results most of the times (i.e. 79% of realizations). However, this result simply indicates that one correctly predicts that observed rates fare badly in estimating the underlying risk (recall Table 2). This approach does not perform as well for least frequent cancers, such as cervix cancer, where the lack of reliability of rates is not captured by the analysis. Because semivariogram estimation is very sensitive to extreme unreliable rates, the sill of the semivariogram model for raw rates greatly varies among the 100 realizations, leading to a wide range of MSSR values: 0.22–13. This lack of robustness prevents the identification of over or under-estimation of the prediction errors in actual applications. Poisson kriging is much more robust with respect to the type of cancer and leads to similar average MSSR for lung and cervix. The range of MSSR for cervix is also much narrower (0.75–4.36), which indicates a better quantification of the magnitude of prediction errors by the kriging variance. The fact that the best results overall are obtained by far for ATP kriging with the true semivariogram model (e.g., the range of cervix MSSR is 0.75–1.57) suggests that future improvements of the method should focus on the estimation of the point-support semivariogram model.

Another way to use the prediction variance is to build at each grid node $\mathbf{u}_s$ the conditional cumulative distribution function (ccdf) of the unknown risk value. Under the assumption of normality of the prediction errors, the ccdf is fully characterized by its mean and variance which are the risk estimate, $r_P^{(l)}(\mathbf{u}_s)$, and the prediction variance, $[\sigma_P^{(l)}(\mathbf{u}_s)]^2$. The probability that the risk variable does not exceed any specific threshold r at $\mathbf{u}_s$ is then:

$$
\begin{aligned}
F_R^{(l)}(\mathbf{u}_s; r \mid (\text{Info})) &= \text{Prob}\{R(\mathbf{u}_s) \le r \mid (\text{Info})\} \\
&= G\left[\left(r - r_P^{(l)}(\mathbf{u}_s) \right) / \sigma_P^{(l)}(\mathbf{u}_s) \right]
\end{aligned}
\qquad (\text{Equation 36})
$$

where $G(\cdot)$ is the cumulative distribution function of the standard normal distribution. The ccdf allows the computation of symmetric *p*-probability intervals (PI) centred on the median; for example, the 0.5-PI is bounded by the lower and upper quartiles $[F_R^{-1}(\mathbf{u}_s; 0.25 \mid (\text{Info})), F_R^{-1}(\mathbf{u}_s; 0.75 \mid (\text{Info}))]$. A correct modeling of local uncertainty would entail that there is a 0.5 probability that the actual risk value at $\mathbf{u}_s$ falls into that interval or, equivalently, that over the study area 50% of the 0.5-PI include the true value. Since the true risk maps are known for our simulation studies, the fraction of true values falling into any given *p*-PI, denoted $\bar{\zeta}^{(l)}(p)$, can be computed easily and compared to the expected fraction *p*. Following Deutsch [32], the agreement between estimated and theoretical fractions is quantified using the following "goodness" statistic:

$$G^{(l)} = 1 - \frac{1}{K'} \sum_{k=1}^{K'} w(p_k) \mid \bar{\zeta}^{(l)}(p_k) - p_k \mid \quad \text{with} \quad 0 \le G^{(l)} \le 1 \qquad (\text{Equation 37})$$

where $w(p_k) = 1$ if $\bar{\zeta}^{(l)}(p_k) > p_k$, and 2 otherwise. The weights *w* penalize the situation where the fraction of true values falling into the *p*-PI is smaller than expected. K' represents the discretization level of the computation; for example, the ccdf percentiles are used as PI boundaries when K' = 50. The goodness statistic confirms the ranking of methods obtained for MSSR; see Table 4 (bottom rows). Point kriging of EBS estimates yields the worst results, while Poisson kriging with the true semivariogram model outperforms other methods. Poisson kriging with the deconvoluted semivariogram model is the 2nd best for

Table 4: Performance comparison of alternative kriging estimators: mean square standardized residual and goodness of uncertainty models.

Estimators	Lung cancer		Cervix cancer	
MSSR	Average	% best result	Average	% best result
Point kriging of raw rates	**1.707**	79	2.176	41
Point kriging of global EBS	5.941	0	4.566	0
Point kriging of local EBS	6.310	0	6.599	0
ATP Poisson kriging	1.827	21	**1.776**	59
ATP Poisson kriging (true $\gamma_R(\mathbf{h})$)	1.081		1.086	
GOODNESS STATISTIC				
Point kriging of raw rates	**0.918**	78	0.868	27
Point kriging of global EBS	0.602	0	0.731	3
Point kriging of local EBS	0.580	0	0.616	0
ATP Poisson kriging	0.862	22	**0.916**	70
ATP Poisson kriging (true $\gamma_R(\mathbf{h})$)	0.981		0.959	

Results obtained on average over 100 realizations generated for Regions 1 and 2. Poisson kriging was conducted with the semivariogram model derived through deconvolution or inferred directly from the simulated grid risk values (true point-support model $\gamma_R(\mathbf{h})$). Bold numbers refer to best performances (i.e. criteria closer to 1) outside the ideal case where the true semivariogram of risk is known. The second column gives the percentage of realizations where the particular method (except ATP kriging with true $\gamma_R(\mathbf{h})$) yields the best results.

the less frequent cervix cancer. Point kriging of raw rates performs better for cancer with high incidence since the corresponding mortality rates are less unstable.

Conclusion

Capitalizing on recent work in the area of change of support and semivariogram deconvolution, this paper generalized the Poisson kriging approach in order to incorporate not only the shape and size of geographical units, but also the spatial repartition of population within those units in the filtering of cancer mortality maps. The new formulation can accommodate any measurement or prediction support, enabling the estimation of mortality risk both at the level of administrative units (area-to-area kriging) or at the nodes of a fine grid discretizing the area of interest (area-to-point kriging). Unlike the empirical geostatistical methods recently proposed for the creation of isopleth maps from areal cancer rates, the present methodology: 1) conducts the filtering and mapping in a single step, 2) ensures the coherence of the prediction, that is the population-weighted average of kriged risks within each geographical unit equals the risk for this unit, and 3) provides a measure of uncertainty (i.e. kriging variance) that accounts for the spatial pattern of mortality risk, the geometry of administrative units and the spatial repartition of the population at risk.

The analysis of age-adjusted lung and cervix cancer mortality rates illustrated some key features of area-to-point (ATP) Poisson kriging relatively to point kriging of raw rates or empirical Bayesian smoothed rates. For both cancers, the coherence constraint implicit to ATP kriging attenuates the smoothness of the kriged maps while the incorporation of the high-resolution population map enhances the impact of low or high rates recorded in the vicinity of urban areas. Because point kriging of areal data arbitrarily assigns the entire county population to the location of its centroid, information about the size and shape of the geographical unit is lost. Point kriging risk maps can either display unrealistically large variability if based on unreliable raw mortality rates or can be over-smoothed if rates are first filtered using an empirical Bayes approach. Differences between prediction methods are even more pronounced with respect to the prediction variance. The point kriging variance is mainly influenced by the location of county centroids: low values simply indicate the presence of clusters of small counties, while large values are observed close to the edges of the study area. A unique feature of the ATP kriging variance is its incorporation of population sizes, leading to lower prediction variance around urban centres. Although the use of population-weighted versus geographic centroids improved the point kriging predictions (results not shown), ATP kriging yields the most accurate predictions for both cancers. Simulation studies also showed that the ATP kriging variance provides a better quantification of the magnitude of prediction errors. The benefit of the proposed approach is the largest for cervix cancer measured in Region 2 because of the combined effect of the low reliability of mortality rates and the wide range of county shapes and sizes.

In earlier papers on the use of Poisson kriging for filtering noise in cancer mortality maps [15], a point kriging approach was implemented whereby only county geographic centroids were considered for semivariogram estimation and kriging. Besides prohibiting the disaggregation of areal data, point Poisson kriging overestimates the prediction variance associated with the areal estimates. The trade-off costs for the implementation of the new form of Poisson kriging are: 1) the inference of the point-support semivariogram from the semivariogram of areal data, 2) the CPU time associated with the computation of area-to-area or area-to-point covariance terms through the fine discretization of geographical units. This paper introduced an innovative deconvolution procedure that allows such an inference in presence of irregular area and heterogeneous population density. Simulation studies demonstrated the convergence of the iterative approach that yields theoretically regularized semivariogram models that are very close to the model inferred from areal data. Although the deconvoluted semivariogram models depart somewhat from the true point-support models, the estimation of risk values appeared to be fairly robust with respect of misspecifications of the semivariogram model. Better prediction of the semivariogram sill would however benefit the quality of the models of uncertainty provided by Poisson kriging. Computational time can be substantially reduced by the use of coarser discretizing grids (e.g. 10 km spacing instead of 5 km). Sensitivity analysis [10] indicates that the minor changes in the deconvoluted model caused by the use of those coarser grids bear little consequences on the kriging predictions.

In their common implementation, full Bayes models fail to account for the size and shape of administrative units, as well as the distribution of the population at risk within those units [17,18]. Simplistic spatial models, such as the popular conditional auto-regressive model, prohibit any change of support and the isopleth mapping of risk values. Therefore, ATP Poisson kriging could only be compared to point kriging of raw rates or empirical Bayesian smoothed rates. Although the latter was introduced as a method for "exploratory" disease mapping [3], it is based on kriging and semivariogram modelling and represents a valid yardstick for the comparison study. When conducting spatial interpolation, one should never forget that every measurement relates to a non-zero finite sample volume. Whereas it is customary in earth science applications to assimilate that measurement support to a point, such a practice becomes a crude simplification for lattice data. Furthermore, treating areal data as point data, then using them to conduct kriging within their own areal support (i.e. disaggregation), is meaningless.

This paper presents a coherent spatial methodology that allows both the filtering of the noise caused by the small number problem and the creation of isopleth maps of mortality risk, thereby avoiding the visual bias associated with the interpretation of choropleth cancer mortality maps. By enabling the estimation of mortality risk over any given support, the approach should facilitate the analysis of relationships between health data and putative covariates (i.e. environmental, socio-economic, behavioral or demographic factors) that are typically measured over different spatial supports. These covariates could also be used directly as secondary information in area-to-point kriging, leading to more detailed risk maps at finer scale [12].

Competing interests

Acknowledgements

References

1. Wakefield J: **A critique of statistical aspects of ecological studies in spatial epidemiology.** *Environmental and Ecological Statistics* 2004, **11:**31-54.
2. Croner CM, De Cola L: **Visualization of Disease Surveillance Data with Geostatistics.** *Presented at UNECE (United Nations Economic Commission for Europe) work session on methodological issues involving integration of statistics and geography: September 2001; Tallinn* [http://www.unece.org/stats/documents/2001/09/gis/25.e.pdf].
3. Berke O: **Exploratory disease mapping: kriging the spatial risk function from regional count data.** *International Journal of Health Geographics* 2004, **3:**18.
4. Oliver MA, Webster R, Lajaunie C, Muir KR, Parkes SE, Cameron AH, Stevens MCG, Mann JR: **Binomial cokriging for estimating and mapping the risk of childhood cancer.** *IMA Journal of Mathematics Applied in Medicine and Biology* 1998, **15:**279-297.
5. Ali M, Goovaerts P, Nazia N, Haq MZ, Yunus M, Emch M: **Application of Poisson kriging to the mapping of cholera and dysentery incidence in an endemic area of Bangladesh.** *International Journal of Health Geographics* 2006, **5:**45.
6. Kelsall J, Wakefield J: **Modeling spatial variation in disease risk: a geostatistical approach.** *Journal of the American Statistical Association* 2002, **97(459):**692-701.
7. Choi K-M, Serre ML, Christakos G: **Efficient mapping of California mortality fields at different spatial scales.** *Journal of Exposure Analysis and Environmental Epidemiology* 2003, **13:**120-133.
8. Journel AG, Huijbregts CJ: *Mining Geostatistics* New York: Academic Press; 1978.
9. Goovaerts P: *Geostatistics for Natural Resources Evaluation* New York: Oxford University Press; 1997.
10. Goovaerts P: **Kriging and semivariogram deconvolution in presence of irregular geographical units.** *Mathematical Geology* 2007. in review
11. Gotway CA, Young LJ: **Combining incompatible spatial data.** *Journal of the American Statistical Association* 2002, **97:**632-648.

12. Gotway CA, Young LJ: **A geostatistical approach to linking geographically-aggregated data from different sources.** In *Technical report # 2004-012* Department of Statistics, University of Florida; 2004.

13. Gotway CA, Young LJ: **Change of support: an inter-disciplinary challenge.** In *geoENV V – Geostatistics for Environmental Applications* Edited by: Renard Ph, Demougeot-Renard H, Froidevaux R. The Netherlands, Springer-Verlag; 2005:1-13.

14. Kyriakidis P: **A geostatistical framework for area-to-point spatial interpolation.** *Geographical Analysis* 2004, **36(3)**:259-289.

15. Goovaerts P: **Geostatistical analysis of disease data: estimation of cancer mortality risk from empirical frequencies using Poisson kriging.** *International Journal of Health Geographics* 2005, **4**:31.

16. Goovaerts P: **Geostatistical analysis of disease data: visualization and propagation of spatial uncertainty in cancer mortality risk using Poisson kriging and p-field simulation.** *International Journal of Health Geographics* 2006, **5**:7.

17. Best N, Richardson S, Thomson A: **A comparison of Bayesian spatial models for disease mapping.** *Statistical Methods in Medical Research* 2005, **14**:35-59.

18. Besag J, York J, Mollie A: **Bayesian image restoration with two applications in spatial statistics.** *Annals of the Institute of Statistical Mathematics* 1991, **43**:1-59.

19. Johnson GD: **Small area mapping of prostate cancer incidence in New York State (USA) using fully Bayesian hierarchical modelling.** *International Journal of Health Geographics* 2004, **3**:29.

20. Best NG, Arnold RA, Thomas A, Waller LA, Conlon EM: **Bayesian models for spatially correlated disease and exposure data.** In *Bayesian Statistics 6* Edited by: Bernardo JM, Berger JO, Dawid AP, Smith AFM. Oxford, UK, Oxford University Press; 1999:131-156.

21. Monestiez P, Dubroca L, Bonnin E, Durbec JP, Guinet C: **Comparison of model based geostatistical methods in ecology: application to fin whale spatial distribution in northwestern Mediterranean Sea.** In *Geostatistics Banff 2004 Volume 2*. Edited by: Leuangthong O, Deutsch CV. Dordrecht, The Netherlands, Kluwer Academic Publishers; 2005:777-786.

22. Monestiez P, Dubroca L, Bonnin E, Durbec JP, Guinet C: **Geostatistical modelling of spatial distribution of Balenoptera physalus in the northwestern Mediterranean Sea from sparse count data and heterogeneous observation efforts.** *Ecological Modelling* 2006, **193**:615-628.

23. Woodward P: **BugsXLA: Bayes for the common man.** *Journal of Statistical Software* 2005, **14**:5.

24. Pickle LW, Mungiole M, Jones GK, White AA: **Exploring spatial patterns of mortality: the new Atlas of United States mortality.** *Statistics in Medicine* 1999, **18**:3211-3220.

25. Grauman DJ, Tarone RE, Devesa SS, Fraumeni JF Jr: **Alternate ranging methods for cancer mortality maps.** *Journal of the National Cancer Institute* 2000, **92(7)**:534-543.

26. Brewer CA, Pickle L: **Evaluation of methods for classifying epidemiological data on choropleth maps in series.** *Annals of the Association of American Geographers* 2002, **92(4)**:662-681.

27. Waller LA, Gotway CA: *Applied Spatial Statistics for Public Health Data* New Jersey: John Wiley and Sons; 2004.

28. Schabenberger O, Gotway CA: *Statistical Methods for Spatial Data Analysis* New York: Chapman & Hall; 2005.

29. Marshall RJ: **Mapping disease and mortality rates using empirical Bayes estimators.** *Applied Statistics* 1991, **40(2)**:283-294.

30. Isaaks EH, Srivastava RM: *An Introduction to Applied Geostatistics* New York: Oxford University Press; 1989.

31. Wackernagel H: *Multivariate Geostatistics, 2nd completely revised edition* Berlin: Springer; 1998.

32. Deutsch CV: **Direct assessment of local accuracy and precision.** In *Geostatistics Wollongong '96 Volume 1*. Edited by: Baafi EY, Schofield NA. Dordrecht, The Netherlands, Kluwer Academic Publishers; 1997:115-125.