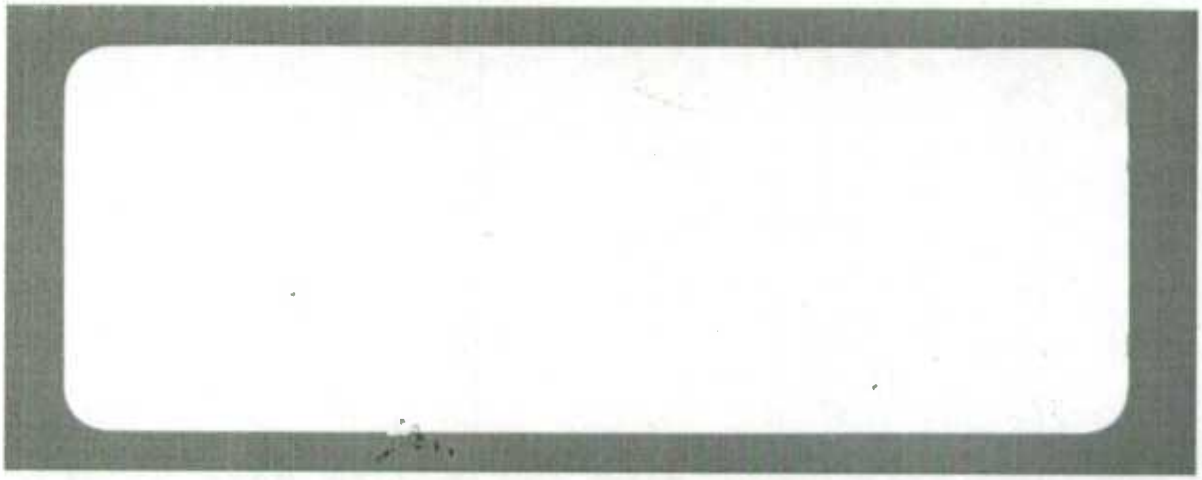
11-619E
no. 2007-007

c.3

Methodology Branch

Direction de la méthodologie



Household Survey Methods Division

Division des méthodes d'enquêtes auprès des ménages

Statistics Canada Statistique Canada

Canada

WORKING PAPER
METHODOLOGY BRANCH

Hierarchical Bayes Small Area Estimation for the Canadian Community Health Survey

HSMD – 2007-007E

Qian M. Zhou and Yong You

Household Survey Methods Division
Statistics Canada

October 2007

The work presented in this paper is the responsibility of the author and does not necessarily represent the views or policies of Statistics Canada

Hierarchical Bayes Small Area Estimation for the Canadian Community Health Survey

Qian M. Zhou and Yong You[1]
University of Waterloo and Statistics Canada

ABSTRACT

In this paper we consider small area health estimation for the Canadian Community Health Survey (CCHS) using a hierarchical Bayes approach. We use cross-sectional area level models including the Fay-Herriot model (Fay and Herriot, 1979) and the model of You and Chapman (You and Chapman, 2006) for sampling variance modeling. In particular, we evaluate different spatial linking models and present a hierarchical Bayes spatial model for analysis of the health data. The proposed hierarchical Bayes spatial model extends the Fay-Herriot model by capturing both the geographically unstructured heterogeneity and spatial correlation effects among areas for local smoothing. The proposed models are implemented using Gibbs sampling approach for fully Bayesian inference. We apply the proposed models to the analysis of Cycle 1.1 of CCHS data and make comparison among the HB model-based estimates and direct design-based estimates. Our results have shown that the HB model-based estimates perform much better than the direct estimates in terms of CV reduction. In addition, the proposed area level spatial models have smaller CVs than the Fay-Herriot model, particularly for the areas with three or more neighbours. Our model has shown that more neighbouring areas can offer more information in the spatial models, and therefore can lead to greater CV reduction over the Fay-Herriot model.

Key Words: Cross-sectional model, Disease rate, Gibbs sampling, Sampling variance, Small area, Spatial model.

[1] Corresponding author: Yong You, yongyou@statcan.ca.

L'estimation bayésienne hiérarchique pour de petits domaines dans l'Enquête sur la santé dans les collectivités canadiennes

Qian M. Zhou and Yong You[1]
University of Waterloo and Statistique Canada

Résumé

Dans cet article, nous examinons les estimations sur la santé pour de petits domaines dans l'Enquête sur la santé dans les collectivités canadiennes (ESCC) en utilisant l'approche bayésienne hiérarchique. Nous utilisons les modèles transversaux au niveau des domaines dont le modèle de Fay-Herriot (Fay et Herriot, 1979) ainsi que le modèle de You et Chapman (You et Chapman, 2006) pour la modélisation de la variance d'échantillonnage. Nous évaluons en particulier différents modèles de couplage spatial et présentons des modèles spatiaux bayésiens pour l'analyse des données sur la santé. Le modèle spatial bayésien hiérarchique que nous proposons élargit le modèle de Fay-Herriot en saisissant les effets de l'hétérogénéité non structurée géographiquement et de la corrélation spatiale entre les domaines en vue d'un lissage local. Les modèles proposés sont mis en œuvre à l'aide de la méthode d'échantillonnage de Gibbs pour obtenir une inférence entièrement bayésienne. Nous appliquons les modèles proposés à l'analyse du Cycle 1.1 des données de l'ESCC et comparons les estimations fondées sur le modèle HB aux estimations directes fondées sur le plan. Nos résultats ont montré que les estimations fondées sur le modèle HB fonctionnent mieux que les estimations directes en ce qui a trait à la réduction du coefficient de variation (c.v.). De plus, les modèles spatiaux au niveau des domaines qui sont proposés ont des c.v. plus faibles que le modèle de Fay-Herriot, particulièrement pour les domaines avec trois voisins ou plus. Notre modèle a permis de démontrer que davantage de domaines voisins peuvent enrichir l'information des modèles spatiaux et ainsi mener à une plus grande réduction des c.v. que le modèle de Fay-Herriot.

MOTS CLÉS : modèle transversal, taux de maladie, échantillonnage de Gibbs, variance d'échantillonnage, petit domaine, modèle spatial.

[1] Corresponding author: Yong You, yongyou@statcan.ca.

1. INTRODUCTION

The Canadian Community Health Survey (CCHS) is a federal survey conducted by Statistics Canada. The primary objective of CCHS is to provide timely and reliable estimates of health determinants, health status and health system utilization across Canada. It is a cross-sectional survey which operates on a two-year collection cycle. The first year of the survey cycle "x.1" targets individuals aged 12 or older who are living in private dwellings, and it is a general population health survey of a large sample (130,000 persons) designed to provide reliable estimates at the health region, provincial and national levels. The second year of the survey cycle "x.2" has a smaller sample (30,000 persons) allocated based on provincial sample buy-ins and is designed to provide provincial and national level results on specific focused health topics. Although national and provincial estimates are very important, there is an increasing demand for health data at lower levels of geography voiced by a number of provinces including British Columbia, Prince Edward Island, Quebec and others. Cycle 1.1 of the CCHS collected data corresponding to 136 health regions in the 10 provinces and three territories. It primarily used two sampling frame. The first one, used as the primary frame, was based on the area frame designed for the Canadian Labour Force Survey, and within the area frame, a multistage stratified cluster design was used to sample dwellings. The second frame consists of a list of telephone numbers. Random digit dialing methodology is used in some of the health regions for cost reasons. The phone numbers are selected using simple random sampling approach. More details of the design are provided in Béland (2002) and Hidiroglou, M. A., Singh, A. and Hamel (2007). In this paper, we are interested in estimating the disease rate for local health regions within provinces, in particular, the disease rate for the 20 health regions in BC province, from the data collected in Cycle 1.1.

Direct estimates, based only on the domain-specific sample data, usually provide reliable estimates of the parameter of interest for large areas such as provinces and nations. However, due to cost or other reasons, it is seldom possible to have a large enough overall sample size to support adequate direct estimates for all the smaller areas of interest. In particular, for small areas such as local health regions and age-sex domains, direct estimators are likely to yield large standard errors. It is necessary to use indirect estimates that borrow strength by using values of the variable of interest from related areas, thus increasing the "effective" sample sizes. These values are brought into the estimation process

through an explicit model that provides a link to related areas through the use of supplementary information such as census counts or administrative records; see Rao (2003) for more discussion on model-based methods. There are two broad classifications for these models: area level models and unit level models. Area level models are based on area direct survey estimators and unit level models are based on individual observations in areas. This paper we focus on area level models.

Area level models such as the Fay-Herriot model (Fay and Herriot, 1979) have been widely used to obtain reliable model-based estimators for small areas. However, in the Fay-Herriot model, several strong assumptions have been made. For example, the sampling variances are assumed to be known, but this assumption is rarely met in practice. You and Chapman (2006) considered the situation where the sampling variances are unknown and estimated individually by direct estimators. Another strong assumption is that the Fay-Herriot model assumes that the area-specific random effects in the linking model are independent and identically distributed. However, in some applications prior knowledge may indicate that geographically close areas tend to have similar values for variables of interest, indicating the existence of locally spatially structured variation. Thus, it may be more realistic to construct spatial models on the area-specific effects to capture the correlation among them.

The objective of this paper is to obtain reliable model-based estimate for the disease rate at health regions level within provinces. In section 2, we propose several area level models which relax the strong assumptions described above based on the basic Fay – Herriot model by incorporating spatial structure on the area-specific effects, and/or assuming the sampling variances unknown. In section 3, we obtain the Hierarchical Bayes estimators of the parameter of interest and the posterior variances through the Gibbs sampling method. In section 4, through the data analysis of Cycle 1.1 of CCHS, we compare the performance of the direct design-based estimates with the model-based estimates, and moreover, compare the proposed models with the basic Fay-Herriot model to investigate the effects of incorporating spatial structure on the area-specific effects. Finally in section 5, we offer some conclusions and discussions.

2. SMALL AREA ESTIMATION MODELS

2.1. Fay-Herriot Model

Let θ_i, the area parameter of interest, denote the underling rate of certain disease for the ith area or health region, where $i = 1, \cdots, m$, and m is the total number of areas. A basic area level model assumes that the θ_i is related to area-specific auxiliary data $\mathbf{x}_i = (x_{i1}, ..., x_{ip})'$ through a linear model

$$\theta_i = \mathbf{x}_i'\beta + v_i, \tag{1}$$

where $\beta = (\beta_1, ..., \beta_p)'$ is the $p \times 1$ vector of regression coefficients, and the v_i's are area-specific random effects assumed to be independent and identically distributed (iid) with $E(v_i) = 0$ and $var(v_i) = \sigma_v^2$. The assumption of normality may also be included. This model is referred to as a linking model for θ_i.

The basic area level model also assumes that a direct survey estimator y_i (usually design-unbiased) of the parameter of interest θ_i is available whenever the area sample size $n_i \succ 1$. It is customary to assume that

$$y_i = \theta_i + e_i, \quad i = 1, \cdots, m \tag{2}$$

where the e_i's are the sampling error associated with the direct estimator y_i. We also assume that the e_i's are independent normal random variables with mean $E(e_i|\theta_i) = 0$ and sampling variance $var(e_i|\theta_i) = \sigma_i^2$. The model (2) is referred to as a sampling model for the direct survey estimator y_i. Combining these two components (1) and (2) leads to the well-known area level linear mixed model, Fay-Herriot model (Fay & Herriot, 1979)

$$y_i = \mathbf{x}_i'\beta + v_i + e_i, \quad i = 1, \cdots, m \tag{3}$$

In the basic Fay-Herriot model (3), the sampling variance σ_i^2 are usually assumed as known, which is a very strong assumption, but it is impractical in many cases. Generally, we can use direct sampling variance estimates from the survey data, however, these direct estimates are unstable if sample sizes are small. Therefore, in practice, a smoothed estimator of σ_i^2 is used in the model and treated as known. In

You (2006, 2008), equal design effects modeling approach was applied to obtain a smooth estimator of sampling variances. The design effect for the *i*th area may be approximately written as

$$deff_i = \frac{s_i^2}{s_{ri}^2}, \text{ for } i = 1, \cdots, m,$$

where s_i^2 is the unbiased direct estimate of sampling variance based on the complex sampling design, and s_{ri}^2 is the estimate of sampling variance based on the assumption of simple random sampling design. For each area, based on the assumption of a common design effect suggested in You (2006, 2008) and Singh, You and Mantel (2005), a smoothed factor *deff* can be obtained by $deff = \sum_{i=1}^{m} deff_i / m$. Then a smoothed sampling variance estimate $\tilde{\sigma}_i^2$ can be obtained by $\tilde{\sigma}_i^2 = s_{ri}^2 \cdot deff$.

Instead of plugging in the smoothed estimates of sampling variances in the model, alternatively we can model the sampling variance directly. In the paper by Wang and Fuller (2003) and You and Chapman (2006), they assume the sampling variance σ_i^2 unknown and estimate σ_i^2 by an unbiased direct estimator s_i^2, which are independent of the direct survey estimator y_i. They also assume that $d_i s_i^2 \sim \sigma_i^2 \chi_{d_i}^2$, where $d_i = n_i - 1$, and n_i is the sample size for the *i*th area. You and Chapman (2006) considered the full HB approach with the Gibbs sampling method which automatically takes into account the extra uncertainty associated with the estimation of σ_i^2. In this paper, we consider both the smoothing and modeling approaches for the sampling variances.

2.2. Spatial Linking Models

Another strong assumption made in the basic Fay – Herriot model (3) is that the area-specific random effects v_i are iid normal variables capturing geographically unstructured heterogeneity among areas. To incorporate spatially-correlated effects in the model, Gaussian Markov random fields (MRF) models are the most commonly used when "neighboring" areas can be defined. In the class of MRF model, the conditional distribution of area-specific random effect v_i in area *i*, given the values of v_j's in all other areas $j \neq i$, depends only on the values of the neighboring areas. Thus in this model, area-specific random effects have a locally dependent prior probability structure, their joint distribution is

determined (up to a normalizing constant) by these conditional distribution. Following the paper by Besag, York and Mollie (1991), the spatial structure on the area-specific effects v_i can be generally specified by a series of conditional distributions written as

$$v_i \mid v_{-i} \sim \mathrm{N}\left(\rho\sum_{j\neq i}\frac{w_{ij}}{w_{i+}}v_j, \frac{\sigma_v^2}{w_{i+}}\right), \text{ for } i=1,\cdots,m \qquad (4)$$

where v_{-i} represents the values of area-specific random effects in all the areas $j \neq i$, w_{ij} are prescribed non-negative weights, with $w_{ij}=0$ unless i and j are neighboring areas, and $w_{i+}=\sum_{j=1}^{m} w_{ij}$. The common choice is $w_{ij}=1$ if i and j are adjacent areas. Such models are also known as Gaussian conditional autoregressive (CAR) model. The parameter ρ, which takes value between -1 and 1, can be regarded as an autocorrelation parameter that characterizes the overall strength of spatial dependence between areas with nonzero weights. An important advantage of the MRF model (4) is that it is possible to make inference about the overall degree of spatial dependence by estimating ρ. However, interpretation of ρ is not straightforward. Moreover, $\rho=0$ indicates independence between areas, but in this case, the MRF model (4) does not reduce to the independent linking model (1) because the variance is not constant across areas. Clayton and Kaldor (1987) used an alternative parameterization as following:

$$v_i \mid v_{-i} \sim \mathrm{N}\left(\rho\sum_{j\neq i} w_{ij}v_j, \sigma_v^2\right), \text{ for } i=1,\cdots,m \qquad (5)$$

Model (5) leads to a joint distribution of $\mathbf{v}=(v_1,\cdots,v_m)'$ as $\mathrm{MVN}(\mathbf{0},\sigma_v^2\mathbf{B})$, where $\mathbf{B}=\mathbf{I}-\rho\mathbf{W}$, and $\mathbf{I}$ is an identity matrix of dimension m and $\mathbf{W}$ is a $m\times m$ adjacent matrix $\mathbf{W}=(w_{ij})$. To ensure a proper joint distribution, the matrix $\mathbf{B}$ must be positive definite, which requires that ρ lie in the range $(1/\lambda_m, 1/\lambda_1)$, where λ_1 and λ_m are the minimum and maximum eigenvalues of the matrix $\mathbf{W}$. If $\rho=0$, the model (5) reduces to the independent linking model (1). Moreover, the model (5) has an invariant conditional variance, but replaces the weighted average in the model (4) by a weighted sum for the conditional mean. However, it seems inappropriate when areas have different numbers of neighbours. In addition, the range of ρ will be defined differently for different models since the values of λ_1 and λ_m depend on the neighbourhood structure (see Best, Richardson and Thomson, 2005). Thus, an

alternative model is proposed by Besag, York and Mollie (1991), to add another spatially dependent random effect u_i in the linking model (1) as follows:

$$\theta_i = \mathbf{x}_i'\beta + v_i + u_i, \tag{6}$$

where u_i's follow the well known intrinsic conditional autoregressive model (Besag et al., 1991) given as

$$u_i \mid u_{-i} \sim \mathrm{N}\left(\frac{\sum_j w_{ij}u_j}{w_{i+}}, \frac{\sigma_u^2}{w_{i+}}\right), \tag{7}$$

where u_{-i} denotes the values of spatial random effects u_j's in all other areas with $j \neq i$. Model (7) is a special case of MRF model (4) by setting ρ to its maximum value 1. In the model (6), the v_i's represent the geographically unstructured components, and u_i's represent the spatial component of between-area variations. In this way, the degree of overall spatial dependence can be expressed based on the proportion of the total (marginal) variation in the $u_i + v_i$ captured by each component.

In practice, it is often unclear how to choose between an unstructured model (e.g., the basic Fay-Herriot model) and a purely spatially structured model (e.g., intrinsic autoregressive model). For model (6), posterior inference about the spatial dependence is based on the proportion of the total variation in the sum of $u_i + v_i$ captured by each component. However, although the univariate conditional distributions of spatial component (7) are well defined, the corresponding joint distribution is improper (with undefined mean and infinite variance). Moreover, the model (6) has potential identifiability problem where only the sum of the random effects $u_i + v_i$ is well identified by the data (e.g., Best, Richardson and Thomson, 2005).

Alternatively, we can consider another spatial parameterization studied by Leroux, Lei, and Breslow (1999) and MacNab (2003), which avoids the identifiability problem encountered with the MRF model (6). Leroux et. al. (1999) and MacNab (2003) placed the following CAR spatial model on the area-specific random effects $\mathbf{v} = (v_1, \cdots, v_m)'$:

$$\mathbf{v} \sim \mathrm{MVN}\left(\mathbf{0}, \Sigma\left(\sigma_v^2, \lambda\right)\right) \tag{8}$$

$$\Sigma\left(\sigma^2, \lambda\right) = \sigma_v^2 \mathbf{D}^{-1}, \ \mathbf{D} = \lambda\mathbf{R} + (1-\lambda)\mathbf{I} \tag{9}$$

where σ_v^2 is a spatial dispersion parameter and λ is a spatial autocorrelation parameter, $0 \le \lambda \le 1$; **I** is an identity matrix of dimension *m*; **R,** commonly known as the neighbourhood matrix, has *i*th diagonal element equal to the number of neighbors w_{i+} of the area *i*, and the off-diagonal elements in each row equal to -1 if the corresponding areas are neighbors and 0 otherwise. The CAR model (8) - (9) results in the following conditional distribution of v_i:

$$v_i \mid v_{-i} \sim \mathrm{N}\left(\frac{\lambda}{1-\lambda+\lambda w_{i+}} \sum_{j \neq i} w_{ij} v_j, \frac{\sigma_v^2}{1-\lambda+\lambda w_{i+}} \right), \; i = 1, \cdots, m$$

The CAR model (8) - (9) becomes the intrinsic autoregressive model if $\lambda = 1$. On the other hand, if $\lambda = 0$, the CAR model (8) - (9) reduces to the independent linking model (1) which assumes independence on the area-specific effects v_i. It is necessary to point out that the conditional mean and variances of $v_i \mid v_{-i}$ are weighted sums of the corresponding "global smoothing" moments from the basic Fay-Herriot model and "local smoothing" moments from the intrinsic autoregressive model:

$$\mathrm{E}(v_i \mid v_{-i}) = \frac{1-\lambda}{1-\lambda+\lambda w_{i+}} \times 0 + \frac{\lambda w_{i+}}{1-\lambda+\lambda w_{i+}} \times \left(\sum_{j \neq} w_{ij} v_j / w_{i+} \right)$$

$$\mathrm{Var}(v_i \mid v_{-i}) = \frac{1-\lambda}{1-\lambda+\lambda w_{i+}} \times \sigma_v^2 + \frac{\lambda w_{i+}}{1-\lambda+\lambda w_{i+}} \times \left(\sigma_v^2 / w_{i+} \right)$$

Thus model (8)-(9) is a balance between the independent linking model (1) and the intrinsic CAR model (7). The spatial correlation parameter λ measures the extent of the spatial effects for "local smoothing" of the neighbouring areas. The modeling structure (9) captures both the unstructured heterogeneity among areas and the spatial correlation effects of the neighbouring area.

3. HIERARCHICAL BAYES INFERENCE

In order to estimate θ_i, the parameter of interest, we apply a hierarchical Bayes (HB) approach using the Gibbs sampling method. Compared to other approaches such as EBLUP and empirical Bayes (EB), HB approach is straightforward and the inference for θ_i are "exact" unlike the EB or EBLUP. Moreover, HB approach can deal with complex small area models using Monte Carlo Markov Chain (MCMC) method, which overcomes the computational difficulties of multi-dimensional integrations of posterior quantities to a large extent.

Let $\mathbf{y} = (y_1, \cdots, y_m)'$, $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_m)'$, and $\mathbf{X} = (\mathbf{x}_1, \cdots, \mathbf{x}_m)'$. We first construct two HB models without and with spatial structure under the assumption that the sampling variance σ_i^2 are assumed known and replaced by the smoothed estimate $\tilde{\sigma}_i^2$.

Model 1: Fay-Herriot model

- $y_i \mid \theta_i \sim \mathrm{N}(\theta_i, \sigma_i^2 = \tilde{\sigma}_i^2)$, for $i = 1, \cdots, m$;
- $\theta_i \mid \beta, \sigma_v^2 \sim \mathrm{N}(\mathbf{x}_i'\beta, \sigma_v^2)$, for $i = 1, \cdots, m$;
- Priors for the parameters (β, σ_v^2): $\pi(\beta) \propto 1$; $\pi(\sigma_v^2) \sim \mathrm{IG}(a_0, b_0)$, where a_0, b_0 are chosen to be very small known constants to reflect vague knowledge on σ_v^2. N stands for normal distribution and IG for inverse gamma distribution.

Model 2: Proposed area level CAR model

- $\mathbf{y} \mid \boldsymbol{\theta} \sim \mathrm{MVN}(\boldsymbol{\theta}, \mathbf{E})$, where $\mathbf{E}$ is a diagonal matrix with the ith diagonal element $\sigma_i^2 = \tilde{\sigma}_i^2$;
- $\boldsymbol{\theta} \mid \beta, \sigma_v^2 \sim \mathrm{MVN}(\mathbf{X}\beta, \sigma_v^2 \mathbf{D}^{-1})$, where $\mathbf{D} = \lambda \mathbf{R} + (1-\lambda)\mathbf{I}$, with $\mathbf{I}$, an identity matrix of dimension m, and $\mathbf{R}$, the neighbourhood matrix;
- Priors for the parameters $(\beta, \lambda, \sigma_v^2)$: $\pi(\beta) \propto 1$; $\pi(\lambda) \sim \mathrm{Uniform}(0,1)$, where $0 \le \lambda \le 1$; $\pi(\sigma_v^2) \sim \mathrm{IG}(a_0, b_0)$, where a_0, b_0 are chosen to be very small known constants. MVN stands for the multivariate normal distribution.

Note that Model 2 reduces to Model 1 when $\lambda = 0$.

We also consider two HB models with the sampling variance σ_i^2 unknown and modeled by the direct unbiased estimator s_i^2.

Model 3: Fay-Herriot model with unknown sampling variances (You and Chapman, 2006)

- $y_i \mid \theta_i, \sigma_i^2 \sim \mathrm{N}(\theta_i, \sigma_i^2)$, for $i = 1, \cdots, m$;
- $d_i s_i^2 \big| \sigma_i^2 \overset{ind}{\sim} \sigma_i^2 \chi_{d_i}^2$, where $d_i = n_i - 1$, for $i = 1, \cdots, m$;

- $\theta_i \mid \beta, \sigma_v^2 \sim \mathrm{N}\left(\mathbf{x}_i'\beta, \sigma_v^2\right)$, for $i = 1, \cdots, m$;
- Priors for the parameters $\left(\beta, \sigma_v^2, \sigma_i^2, i = 1, \cdots, m\right)$: $\pi(\beta) \propto 1$; $\pi(\sigma_v^2) \sim \mathrm{IG}(a_0, b_0)$, $\pi(\sigma_i^2) \sim \mathrm{IG}(a_i, b_i)$ for $i = 1, \cdots, m$, where a_i, b_i $\left(0 \le i \le m\right)$ are chosen to be very small known constants to reflect vague knowledge on σ_i^2 and σ_v^2.

Model 4: Proposed area level CAR model with unknown sampling variances

- $\mathbf{y} \mid \boldsymbol{\theta}, \sigma_1^2, \cdots, \sigma_m^2 \sim \mathrm{MVN}\left(\boldsymbol{\theta}, \mathbf{E}\right)$, where $\mathbf{E}$ is a diagonal matrix with the ith diagonal element σ_i^2;
- $d_i s_i^2 \left| \sigma_i^2 \right. \overset{ind}{\sim} \sigma_i^2 \chi_{d_i}^2$, where $d_i = n_i - 1$, for $i = 1, \cdots, m$;
- $\boldsymbol{\theta} \mid \beta, \sigma_v^2 \sim \mathrm{MVN}(\mathbf{X}\beta, \sigma_v^2 \mathbf{D}^{-1})$, where $\mathbf{D} = \lambda \mathbf{R} + (1 - \lambda)\mathbf{I}$;
- Priors for the parameters $\left(\beta, \lambda, \sigma_v^2, \sigma_i^2, i = 1, \cdots, m\right)$: $\pi(\beta) \propto 1$; $\pi(\lambda) \sim \mathrm{Uniform}(0,1)$, where $0 \le \lambda \le 1$; $\pi(\sigma_v^2) \sim \mathrm{IG}(a_0, b_0)$; $\pi(\sigma_i^2) \sim \mathrm{IG}(a_i, b_i)$ for $i = 1, \cdots, m$, where a_i, b_i $\left(0 \le i \le m\right)$ are chosen to be very small known constants.

Again, note that Model 4 reduces to Model 3 when $\lambda = 0$.

In the HB approach, we use the posterior mean $\mathrm{E}(\theta_i \mid \mathbf{y})$ as a point estimate of θ_i and the posterior variance $\mathrm{Var}(\theta_i \mid \mathbf{y})$ as a measure of variability. We implement Gibbs sampling method (Gelfand and Smith, 1990) by drawing samples $\left\{\theta_1^{(k)}, \cdots, \theta_m^{(k)}, \beta^{(k)}, \sigma_v^{2(k)}\right\}$ from the joint posterior distribution $p\left(\theta_1, \cdots, \theta_m, \beta, \sigma_v^2\right)$ for Model 1, or drawing samples $\left\{\theta_1^{(k)}, \cdots, \theta_m^{(k)}, \beta^{(k)}, \lambda^{(k)}, \sigma_v^{2(k)}\right\}$ from the joint posterior distribution $p\left(\theta_1, \cdots, \theta_m, \beta, \lambda, \sigma_v^2\right)$ for Model 2, or drawing samples $\left\{\theta_1^{(k)}, \cdots, \theta_m^{(k)}, \beta^{(k)}, \sigma_v^{2(k)}, \sigma_1^{2(k)}, \cdots, \sigma_m^{2(k)}\right\}$ from the joint posterior distribution $p\left(\theta_1, \cdots, \theta_m, \beta, \sigma_v^2, \sigma_1^2, \cdots, \sigma_m^2\right)$ for Model 3, or drawing samples $\left\{\theta_1^{(k)}, \cdots, \theta_m^{(k)}, \beta^{(k)}, \lambda^{(k)}, \sigma_v^{2(k)}, \sigma_1^{2(k)}, \cdots, \sigma_m^{2(k)}\right\}$ from the joint posterior distribution $p\left(\theta_1, \cdots, \theta_m, \beta, \lambda, \sigma_v^2, \sigma_1^2, \cdots, \sigma_m^2\right)$ for Model 4.

For Model 1, the full conditional distributions of $\left(\theta_1, \cdots, \theta_m, \beta, \sigma_v^2\right)$ for the Gibbs sampler are:

- $\left[\theta_i \mid y_i, \beta, \sigma_v^2\right] \sim \mathrm{N}\left[\gamma_i y_i + (1 - \gamma_i)\mathbf{x}_i'\beta,\ \tilde{\sigma}_i^2 \gamma_i\right]$, where $\gamma_i = \dfrac{\sigma_v^2}{\sigma_v^2 + \tilde{\sigma}_i^2}$, for $i = 1, \cdots, m$;

- $\left[\beta \mid \boldsymbol{\theta}, \sigma_v^2\right] \sim \mathrm{N}\left[\left(\sum_{i=1}^{m} \mathbf{x}_i \mathbf{x}'\right)^{-1}\left(\sum_{i=1}^{m} \mathbf{x}_i \theta_i\right), \sigma_v^2\left(\sum_{i=1}^{m} \mathbf{x}_i \mathbf{x}_i'\right)^{-1}\right]$;
- $\left[\sigma_v^2 \mid \boldsymbol{\theta}, \beta\right] \sim \mathrm{IG}\left[a_0 + \frac{1}{2}m,\ b_0 + \frac{1}{2}\sum_{i=1}^{m}(\theta_i - \mathbf{x}_i'\beta)^2\right]$.

It is straightforward to draw samples from these full conditional distributions.

For Model 2, the full conditional distributions of $\left(\boldsymbol{\theta}, \beta, \lambda, \sigma_v^2\right)$ for the Gibbs sampler are:

- $\left[\boldsymbol{\theta} \mid \mathbf{y}, \beta, \lambda, \sigma_v^2\right] \sim \mathrm{MVN}(\boldsymbol{\Lambda}\mathbf{y} + (\mathbf{I} - \boldsymbol{\Lambda})\mathbf{X}\beta, \boldsymbol{\Lambda}\mathbf{E})$, where $\boldsymbol{\Lambda} = \left(\mathbf{E}^{-1} + \mathbf{D}/\sigma_v^2\right)^{-1}\mathbf{E}^{-1}$ with $\mathbf{E} = \mathrm{diag}\left\{\tilde{\sigma}_1^2, \cdots, \tilde{\sigma}_m^2\right\}$ and $\mathbf{D} = \lambda\mathbf{R} + (1-\lambda)\mathbf{I}$;
- $\left[\beta \mid \boldsymbol{\theta}, \lambda, \sigma_v^2\right] \sim \mathrm{MV\,N}\left[(\mathbf{X}'\mathbf{D}\mathbf{X})^{-1}\mathbf{X}'\mathbf{D}\boldsymbol{\theta}, \sigma_v^2(\mathbf{X}'\mathbf{D}\mathbf{X})^{-1}\right]$;
- $\left[\lambda \mid \boldsymbol{\theta}, \beta, \sigma_v^2\right] \propto \left|\left[\lambda\mathbf{R} + (1-\lambda)\mathbf{I}\right]^{-1}\right|^{\frac{1}{2}} \times \exp\left\{-\frac{1}{2\sigma_v^2}(\boldsymbol{\theta} - X\beta)'\left[\lambda\mathbf{R} + (1-\lambda)\mathbf{I}\right](\boldsymbol{\theta} - X\beta)\right\}$
- $\left[\sigma_v^2 \mid \boldsymbol{\theta}, \beta, \lambda\right] \sim \mathrm{IG}\left[a_0 + \frac{m}{2}, b_0 + \frac{1}{2}(\boldsymbol{\theta} - X\beta)'\mathbf{D}(\boldsymbol{\theta} - X\beta)\right]$.

The distributions of $\boldsymbol{\theta}$, β and σ_v^2 are standard multivariate normal or inverse gamma distributions that can be easily sampled. However, the conditional distribution of λ does not have a closed form. We use the Metropolis-Hastings algorithm within the Gibbs sampler (Chip and Greenberg, 1995) to update λ. The full conditional distribution of λ in the Gibbs sampler can be written as

$$\left[\lambda \mid \boldsymbol{\theta}, \beta, \sigma_v^2\right] \propto \mathrm{h}(\lambda)\mathrm{f}(\lambda)$$

where $\mathrm{f}(\lambda)$ is a density function of the uniform distribution Uniform(0,1) given as

$$\mathrm{f}(\lambda) \propto 1, \text{ where } 0 \le \lambda \le 1$$

and $\mathrm{h}(\lambda)$ is a function given by

$$\mathrm{h}(\lambda) \propto \left|\left[\lambda\mathbf{R} + (1-\lambda)\mathbf{I}\right]^{-1}\right|^{\frac{1}{2}} \times \exp\left\{-\frac{1}{2\sigma_v^2}(\boldsymbol{\theta} - X\beta)'\left[\lambda\mathbf{R} + (1-\lambda)\mathbf{I}\right](\boldsymbol{\theta} - X\beta)\right\}.$$

We use $\mathrm{f}(\lambda)$ as the "candidate" generating density function in the Metropolis-Hastings updating step. To update λ from the current values of $\left(\boldsymbol{\theta}^{(k)}, \beta^{(k)}, \sigma_v^{2(k)}\right)$, we proceed as follows:

(1) Draw λ^* from a uniform distribution Uniform(0,1)

(2) Compute the acceptance probability

$\alpha(\lambda^*, \lambda^{(k)}) = \min\{h(\lambda^*)/h(\lambda^{(k)}), 1\}$

(3) Generate u from a uniform distribution Uniform(0,1), if $u < \alpha(\lambda^*, \lambda^{(k)})$, then this candidate value λ^* is accepted, i.e., $\lambda^{(k+1)} = \lambda^*$; otherwise λ^* is rejected, and set $\lambda^{(k+1)} = \lambda^{(k)}$.

For Model 3, the full conditional distributions of $(\theta_1, \cdots, \theta_m, \beta, \sigma_v^2, \sigma_1^2, \cdots, \sigma_m^2)$ for the Gibbs sampler are:

- $[\theta_i \mid y_i, \beta, \sigma_i^2, \sigma_v^2] \sim \mathrm{N}[\gamma_i y_i + (1-\gamma_i)\mathbf{x}_i'\beta, \sigma_i^2\gamma_i]$, where $\gamma_i = \dfrac{\sigma_v^2}{\sigma_v^2 + \sigma_i^2}$, for $i = 1, \cdots, m$;
- $[\beta \mid \boldsymbol{\theta}, \sigma_v^2] \propto \mathrm{N}\left[\left(\sum_{i=1}^m \mathbf{x}_i\mathbf{x}_i'\right)^{-1}\left(\sum_{i=1}^m \mathbf{x}_i\theta_i\right), \sigma_v^2\left(\sum_{i=1}^m \mathbf{x}_i\mathbf{x}_i'\right)^{-1}\right]$;
- $[\sigma_i^2 \mid y_i, \theta_i] \sim \mathrm{IG}\left(a_i + \dfrac{d_i+1}{2}, b_i + \dfrac{(y_i-\theta_i)^2 + d_i s_i^2}{2}\right)$, where $d_i = n_i - 1$, for $i = 1, \cdots, m$;
- $[\sigma_v^2 \mid \boldsymbol{\theta}, \beta] \sim \mathrm{IG}\left[a_0 + \dfrac{1}{2}m,\ b_0 + \dfrac{1}{2}\sum_{i=1}^m (\theta_i - \mathbf{x}_i'\beta)^2\right]$.

Similarly to Model 1, it is easy to draw samples from these full conditional distributions.

For Model 4, the full conditional distributions of $(\boldsymbol{\theta}, \beta, \lambda, \sigma_v^2, \sigma_1^2, \cdots, \sigma_m^2)$ for the Gibbs sampler are:

- $[\boldsymbol{\theta} \mid \mathbf{y}, \beta, \lambda, \sigma_v^2, \sigma_1^2, \cdots, \sigma_m^2] \sim \mathrm{MVN}(\boldsymbol{\Lambda}\mathbf{y} + (\mathbf{I} - \boldsymbol{\Lambda})\mathbf{X}\beta, \boldsymbol{\Lambda}\mathbf{E})$, where $\boldsymbol{\Lambda} = (\mathbf{E}^{-1} + \mathbf{D}/\sigma^2)^{-1}\mathbf{E}^{-1}$ with $\mathbf{E} = \mathrm{diag}\{\sigma_1^2, \cdots, \sigma_m^2\}$ and $\mathbf{D} = \lambda\mathbf{R} + (1-\lambda)\mathbf{I}$;
- $[\beta \mid \boldsymbol{\theta}, \lambda, \sigma_v^2] \sim \mathrm{MVN}[(\mathbf{X}'\mathbf{D}\mathbf{X})^{-1}\mathbf{X}'\mathbf{D}\boldsymbol{\theta}, \sigma_v^2(\mathbf{X}'\mathbf{D}\mathbf{X})^{-1}]$;
- $[\lambda \mid \boldsymbol{\theta}, \beta, \sigma_v^2] \propto |[\lambda\mathbf{R} + (1-\lambda)\mathbf{I}]^{-1}|^{\frac{1}{2}} \times \exp\left\{-\dfrac{1}{2\sigma_v^2}(\boldsymbol{\theta} - X\beta)'[\lambda\mathbf{R} + (1-\lambda)\mathbf{I}](\boldsymbol{\theta} - X\beta)\right\}$
- $[\sigma_i^2 \mid y_i, \theta_i] \sim \mathrm{IG}\left(a_i + \dfrac{d_i+1}{2}, b_i + \dfrac{(y_i-\theta_i)^2 + d_i s_i^2}{2}\right)$, where $d_i = n_i - 1$, for $i = 1, \cdots, m$;
- $[\sigma_v^2 \mid \boldsymbol{\theta}, \beta, \lambda] \sim \mathrm{IG}\left[a_0 + \dfrac{m}{2}, b_0 + \dfrac{1}{2}(\boldsymbol{\theta} - X\beta)'\mathbf{D}(\boldsymbol{\theta} - X\beta)\right]$.

Similar to the full conditionals for Model 2, the conditional distribution of λ does not have a closed form. We apply the same procedure to update λ using Metropolis-Hastings algorithm within the Gibbs sampler as in Model 2.

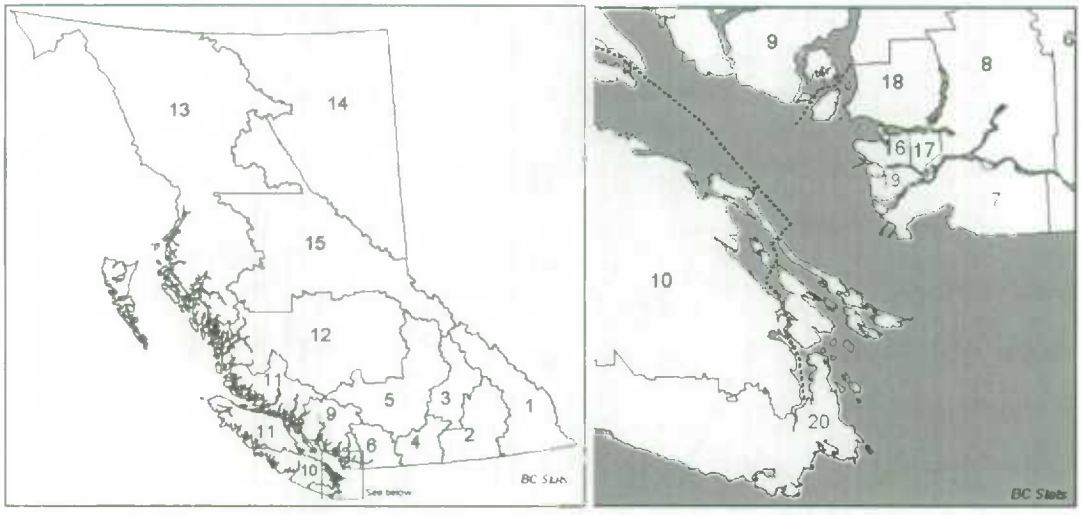
To implement the Gibbs sampling, we use L=5 parallel runs each with a "burn-in" length of B=2000 and Gibbs sampling size of G=5000. For Model 2 and Model 4, in order to reduce the autocorrelation which results from the accept-rejection algorithm in the run, we take every 5^{th} iteration after the "burn-in" period. Therefore, for Model 1 and Model 3, we have n=5000 samples for each run, and for Model 2 and Model 4, n=1000 samples for each run.

In the following section, we apply the proposed four HB models in Section 3 to estimate disease rates for health regions using CCHS Cycle 1.1 data.

4. DATA ANALYSIS

The health region-level survey of Cycle 1.1 consists of common content to meet basic health data requirements on an on-going basis. The questionnaire contains the information about several chronic health conditions, including food allergies, asthma, arthritis or rheumatism, diabetes and etc. In this paper, we are interested in estimating the rate of asthma in the 20 health regions of BC province. Figure 1 shows the map of the 20 health regions in the province of British Columbia.

Figure 1: The map of 20 health regions in the province of British Columbia.



Based on the map in Figure 1, for each health region, we define the corresponding adjacent health regions given in following table:

Table 1: The neighbouring areas defined for the 20 health regions

Health Region Number	Health Region Name	Number of Health Regions in the Neighbouring Area	Health Region Numbers in the Neighbouring Area
1	East Kootenay	3	2, 3, 15
2	West Kootenay-Boundary	3	1, 3, 4
3	North Okanagan	5	1, 2, 4, 5, 15
4	South Okanagan Similameen	4	2, 3, 5, 6
5	Thompson	7	3, 4, 6, 9, 11, 12, 15
6	Fraser Valley	5	4, 5, 7, 8, 9
7	South Fraser Valley	4	6, 8, 17, 19
8	Simon Fraser	5	6, 7, 9, 17, 18
9	Coast Garibaldi	5	5, 6, 8, 11, 18
10	Central Vancouver Island	2	11, 20
11	Upper Island/Central Coast	4	5, 9, 10, 12
12	Cariboo	4	5, 11, 13, 15
13	North West	3	12, 14, 15
14	Peace Liard	2	13, 15
15	Northern Interior	6	1, 3, 5, 12, 13, 14
16	Vancouver	4	17, 18, 19, 20
17	Burnaby	5	7, 8, 16, 18, 19
18	North Shore	4	8, 9, 16, 17
19	Richmond	3	7, 16, 17
20	Capital	2	10, 16

It is necessary to point out that the two health regions "20 Capital" and "16 Vancouver" are not adjacent since they are separated by the ocean. However, due to the intensive connection between these

two areas when considering transportation, economics, tourism and other aspects, it is reasonable to define that they are neighbours as well.

From the survey data of Cycle 1.1, we obtained eight variables for each health region to estimate the rate of asthma as follows: (1) sample size, (2) direct estimate of the number of persons who have asthma, (3) total population size, (4) number of persons who have asthma as one of the symptoms of the chronic disease, (5) number of persons who have asthma as the main symptom of the chronic disease, (6) number of persons who have diabetes as one of the symptoms of the chronic disease, (7) number of persons who have diabetes as the main symptom of the chronic disease, and (8) number of visits to hospitals. For each health region, the direct estimate y_i of the rate of asthma θ_i is obtained as the ratio of number of people having asthma over the corresponding population size, i.e.,

$$y_i = \frac{\text{direct estimate of the number of persons who have asthma in area } i}{\text{total population size}}$$

for $i = 1, \cdots, m$. The six variables 3, 4, 5, 6, 7, and 8 are used as the area-specific auxiliary data $\mathbf{x}_i = (x_{i1}, ..., x_{i6})'$.

In the literatures related to disease mapping (e.g., Mollié, 1996; Maiti, 1998; MacNab 2003), Poisson or Binomial distribution is usually assumed in the sampling model for the direct estimator y_i. However, in our application, the direct estimator y_i is obtained based on the complex sampling design used in the survey. Thus, it is more reasonable to assume normal approximation on the direct estimator y_i.

At first, we present the HB estimates of the rate of asthma under the Model 1 and 2 in which the sampling variances σ_i^2 are assumed to be known. We use the smoothed estimate $\tilde{\sigma}_i^2$ obtained by the smoothing technique in You (2008) described in Section 2. Figure 2 displays the direct estimates and the HB model-based estimates from Model 1 and Model 2 for the 20 health regions in the province of BC. The health regions appear in the x-coordinate ranked by the order of sample size with the largest (South Fraser Valley) on the left and the smallest (Peace Liard) on the right. Model 1 (Fay-Herriot model) and Model 2 (CAR model) give similar point estimates, and both the model-based estimates

lead to moderate smooth estimates compared to the direct estimates. Moreover, the direct estimates and two HB estimates of the disease rate are very close for some health regions with large sample sizes, but for some areas with smaller sample sizes, they differ to some extent.

Figure 2: Comparison of direct and HB model-based estimates
under the Fay-Herriot model 1 and CAR model 2

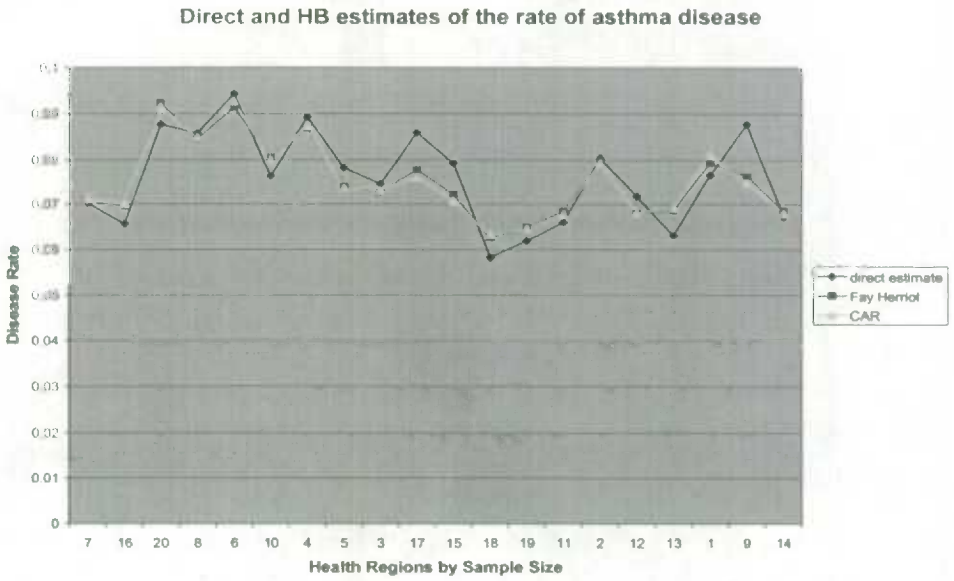


Figure 3 presents the CVs of the direct and two HB model-based estimates with the health regions ordered by the sample sizes from the largest to the smallest. The CVs of HB estimates are obtained by dividing the squared root of the posterior variance by the posterior mean. As expected, the CVs of the direct estimates show a clear tendency of increasing as the sample size decreases, which demonstrates the unreliability of direct estimates in the areas with small sample sizes. However, the two model-based estimates give smoother CVs. Moreover, the two HB model-based estimates exhibit a great improvement over the direct design-based estimates in terms of precision and reliability, that is, smaller CVs. Compared to the direct estimates, the average CV reduction of the HB estimates under Model 1 (Fay-Herriot model) is about 23.97% ranging from 13.45% to 34.48%, and the average reduction of the CVs for the HB estimates under Model 2 (CAR model) is 28.9% ranging from 16.01% to 40.49%.

Figure 3: Comparison of direct and HB CVs under the Fay-Herriot model 1 and CAR model 2 with the health regions sorted by sample size

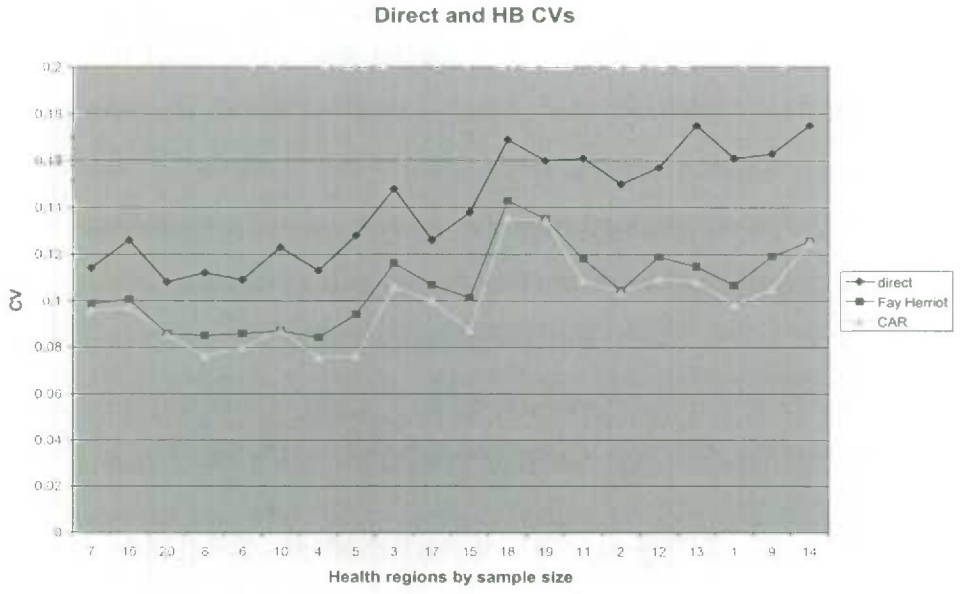


Figure 4 also displays the CVs of the direct and HB estimates, while the health regions are sorted by the number of neighbouring regions from largest to smallest in order to investigate the effects of incorporating the spatial structure in the model. It shows that the HB estimates from the CAR Model 2 has smaller CVs than the estimates from the basic Fay-Herriot model. In addition, the improvement of the CAR model over the Fay-Herriot model is larger in the areas with more neighbours. Howver, these two models give very close CVs in the regions with less adjacent areas. Table 2 lists the reduction of the CVs under the CAR model over the Fay-Herriot model across the health regions with the same number of neighbours. The results in Table 2 present the CV reduction of the CAR model for both cases of known and unknown sampling variances. For example, for known σ_i^2 (smoothed $\tilde{\sigma}_i^2$), for areas with only 2 neighbours, the average CV reduction of CAR model over the Fay-Herriot model is only aroung 0.5%, whereas for areas with 7 neighbours, the average CV reduction for CAR model is as high as around 20%. The numerical results also confirm the clear trend of more CV reduction under the CAR model over the Fay-Herriot model as the number of neighbours increases. Thus, more neighbouring areas provide more information in the spatial structure to improve the precision and reliability of the HB estimates.

Figure 4: Comparison of direct and HB CVs under the Fay-Herriot model 1 and CAR model 2 with the health regions sorted by the number of neighbours

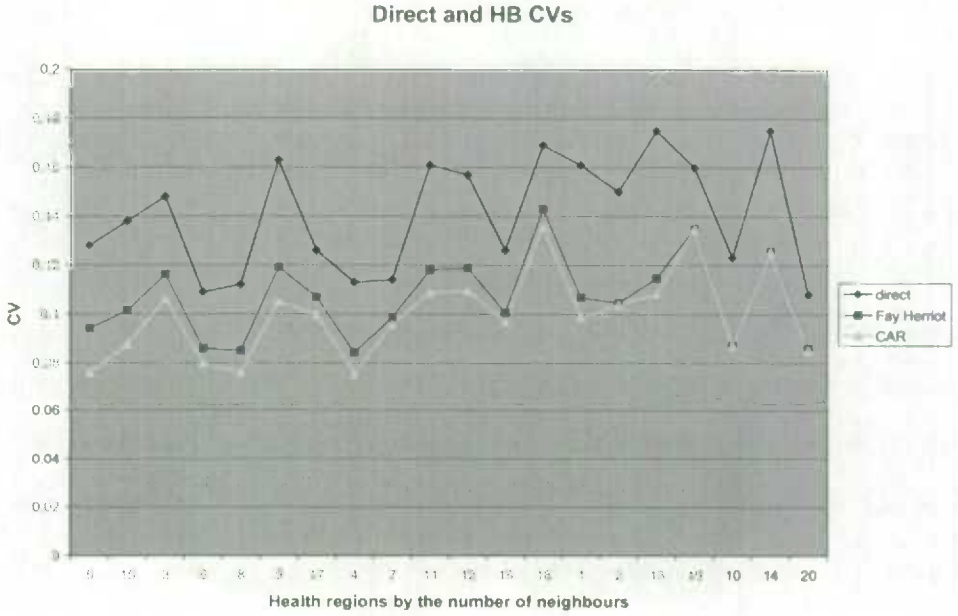


Table 2: The average CV reduction of CAR model over the Fay Herriot model

The number of neighbours	The average CV reduction σ_i^2 known	σ_i^2 unknown
7	**19.2%**	**20.7%**
6	**13.7%**	**11.0%**
5	**8.9%**	**8.7%**
4	**6.3%**	**6.0%**
3	**3.7%**	**3.5%**
2	**0.5%**	**1.8%**

We also obtained point estimates and CVs under Model 3 and Model 4. The comparison of the results under Model 3 and Model 4 has shown similar results with the comparison among the direct estimates and two HB model-based estimates under the Model 1 and 2.

5. CONCLUSION AND DISCUSSION

In this paper we have studied the well-known Fay-Herriot model in which two strong assumptions are made. One is that the sampling variances σ_i^2 are assumed to be known. You (2006, 2008) used the smoothed variances $\tilde{\sigma}_i^2$ obtained by the equal design effects modeling approach, and You and Chapman (2006) instead modeled the sampling variances σ_i^2 directly by the unbiased estimator s_i^2. The other assumption is that the area-specific random effects are assumed independent and identically distributed. Various forms of Gaussian CAR model were proposed in the literature for disease mapping to incorporate spatially-correlated effects. According to the previous work, we propose four HB models which relax these two strong assumptions to investigate the effect of including the geographically structured distribution in the model where the sampling variances σ_i^2 are replaced by the smoothed estimate $\tilde{\sigma}_i^2$ or modeled by the direct estimator s_i^2.

In the data analysis which aims at estimating the rate of asthma for the 20 health regions in the province of British Columbia, the model-based estimates achieve a great improvement over the direct estimates in terms of moderately smoothed point estimates and much smaller CVs. In addition, we find that whenever the sampling variances are assumed to be known or unknown, the proposed area level CAR models have smaller CVs than the Fay-Herriot model which imposes independent area-specific random effects. Moreover, the CV reduction of CAR model over the Fay-Herriot model is greater for the areas with more neighbours.

One possible limitation of our proposed model is that the linking model for the disease rate θ_i is a linear model with normal random effects. Since θ_i takes value between 0 and 1, and it is close to 0 for some rare disease, the linear linking model with normal random effects may lead to negative estimates for θ_i for some small areas in practice if the sampling variances vary substantially. You and Rao (2002) proposed a log-linear linking model for the Fay-Herriot model as the unmatched sampling and linking models as follows:

$y_i = \theta_i + e_i, \quad i = 1, \cdots, m$

$\log(\theta_i) = \mathbf{x}_i' \beta + v_i, \quad i = 1, \cdots, m$

In future work, the proposed CAR models can be extended to the unmatched sampling and linking models with the sampling variance known or unknown. We will also plan to evaluate the estimation effects of different spatial models (e.g., Best, Richardson and Thomson, 2005) as well as the effects of spatial structures. We also plan to study different methods to test the overall fit of the proposed models, and also to assess model fit at the individual area level. For data analysis, we will produce model-based health status estimates based on the proposed models for health regions across Canada and evaluate the possibility of extending the model-based approach to lower level estimates such as age-sex domains within heath regions. One way to do this is to extend the area level models to unit level model with spatial correlation structure between area level variations.

REFERENCES

Béland, Y. (2002). Canadian Community Health Survey Methodological overview. Health report, Statistics Canada, Catalogue no. 82-003-XPE, Vol. 13, No. 3, ISSN 0840-6529.

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

Best, N., Richardson S. and Thomson, A. (2005). A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research*, 14, 35-39.

Chip, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49, 327-335.

Clayton, D. and Kaldor, J. (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*, 43, 359-370.

Fay, R. E. and Herriot, R. A. (1979). Estimation of income for small places: An application of James-Stein procedures to census data, *Journal of the American Statistical Association*, 74, 268-277.

Gelfand, A. E. and Smith, A. F. M. (1990). Gibbs sampling for marginal posterior expectations. *Communications In Statistics – Theory and Methods*, 20, 1747-1766.

Hidiroglou, M. A., Singh, A., Hamel, M. (2007). Some thoughts on small area estimation for the Canadian Community Health Survey (CCHS). Internal report, Statistics Canada.

Leroux, B. G., Lei, X., Breslow, N. (1999). Estimation of disease rates in small areas : a new mixed model for spatial dependence. In *Statistical Models in Epidemiology, the Environment and Clinical Trials*, Halloran ME, Berry D (eds). Springer-Verlag: New York, 135-178.

MacNab, Y. C. (2003). Hierarchial Bayesian spatial modeling of small-area rates of non-rare disease. *Statistics in Medicine*, 22, 1761-1773.

Maiti, T. (1998). Hierarchical Bayes estimation of mortality rates for disease mapping. *Journal of Statistical Planning and Inference*, 69, 339-348.

Mollie, A. (1996). Bayesian mapping of disease. In Gilks, W.R., Richardon, S., Spiegelhalter, D.J., editors. In *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall, 359-379.

Rao, J. N. K. (2003). *Small Area Estimation*. John Wiley & Sons, New York.

Singh, A., You, Y. and Mantel, H. (2005). Use of generalized design effects for variance function modeling in small area estimation from survey data. Presentation at the 2005 Statistical Society of Canada Annual Meeting, Regina, SK.

Wang, J. and Fuller, W. A. (2003). The mean square error of small area predictors constructed with estimated area variances. *Journal of the American Statistical Association*, 98, 716-723.

You, Y. (2006). Model-based small area unemployment rate estimation for the Candian Labour Force Survey. Methodology Branch working paper, HSMD-2006-004E, Statistics Canada.

You, Y. (2008). An integrated modeling approach to unemployment rate estimation for sub-provincial areas of Canada. *Survey Methodology*, 34, to appear.

You, Y. and Chapman, B. (2006). Small area estimation using area level models and estimated sampling variances. *Survey Methodology*, 32, 97-103.

You, Y. and Rao, J. N. K. (2002). Small area estimation using unmatched sampling and linking models. *The Canadian Journal of Statistics*, 20, 3-15.

STATISTICS CANADA LIBRARY
BIBLIOTHÈQUE STATISTIQUE CANADA
1010437462