

Comparing definitions of spatial relations for the analysis of geographic disparities in mortality within a Bayesian mixed-effects framework

Comparação de definições de relações espaciais para a análise das desigualdades geográficas na mortalidade no quadro de efeitos mistos bayesiana

Diego Fernando Rojas-Gualdrón¹

ABSTRACT: *Objective:* To analyze the conceptual and technical differences between three definitions of spatial relations within a Bayesian mixed-effects framework: classical multilevel definition, spatial multiple membership definition and conditional autoregressive definition with an illustration of the estimate of geographic disparities in early neonatal mortality in Colombia, 2011-2014. *Methods:* A registry based cross-sectional study was conducted. Births and early neonatal deaths were obtained from the Colombian vital statistics registry for 2011-2014. Crude and adjusted Bayesian mixed effects regressions were performed for each definition of spatial relation. Model fit statistics, spatial autocorrelation of residuals and estimated mortality rates, geographic disparity measures, relative ratios and relative differences were compared. *Results:* The definition of spatial relations between municipalities based on the conditional autoregressive prior showed the best performance according to both fit statistics and residual spatial pattern analyses. Spatial multiple membership definition had a poor performance. *Conclusion:* Bayesian mixed effects regression with conditional autoregressive prior as an analytical framework may be an important contribution to epidemiological design as an improved alternative to ecological methods in the analyses of geographic disparities of mortality, considering potential ecological bias and spatial model misspecification.

Keywords: Health status disparities. Early neonatal mortality. Gross domestic product. Vital statistics. Spatial regression. Bayesian analysis.

¹Facultad de Medicina, Universidad CES, Medellín, Antioquia, Colombia.

Corresponding author: Diego Fernando Rojas-Gualdrón. Calle 10 A No. 22 – 04,050021 – Medellín, Antioquia, Colombia. E-mail: drojas3@unab.edu.co

Conflict of interests: nothing to declare – **Financial support:** Rojas-Gualdrón receives funding for doctoral studies from COLCIENCIAS.

RESUMO: *Objetivo:* Analisar as diferenças conceituais e técnicas entre três definições de relações espaciais dentro do quadro de efeitos mistos bayesiano: definição multinível clássica, definição de filiação múltipla espacial e definição condicional auto regressivo com uma ilustração da estimativa das disparidades geográficas na mortalidade neonatal precoce na Colômbia, 2011-2014. *Métodos:* Foi realizado um estudo transversal de base do registro. Nascimentos e mortes neonatais precoces foram obtidos a partir do registro de estatísticas vitais Colombiano para o período 2011-2014. Regressões mistas bayesianas brutas e ajustadas foram realizadas para cada definição de relação espacial. As estatísticas de ajuste do modelo, autocorrelação espacial dos resíduos, as estimativas das taxas de mortalidade, as medidas de disparidade geográfica, as relações relativas e as diferenças relativas foram comparadas. *Resultados:* A definição das relações espaciais entre os municípios com base no priori condicional auto regressivo apresentou o melhor desempenho de acordo com as estatísticas de ajuste e a análises de padrão espacial dos resíduos. A definição de filiação múltipla espacial mostrou o mau desempenho. *Conclusão:* A regressão de efeitos mistos bayesiana com priori condicional auto regressivo como quadro analítico pode ser uma contribuição importante para desenho epidemiológico como uma alternativa melhorada aos métodos ecológicos nas análises das desigualdades geográficas, considerando e potencial viés ecológico e má especificação do modelo espacial.

Palavras-chave: Disparidades nos níveis de saúde. Mortalidade neonatal precoce. Produto interno bruto. Estatísticas vitais. Regressão espacial. Análise bayesiana.

INTRODUCTION

In response to growing interest in social and spatial epidemiology studies, which attempt to explain both individual probability of death and between-group disparities within this probability, different methodological approaches for area-level data have been proposed in order to simultaneously model them. Many of these proposals are modifications or extensions of mixed-effects modeling¹. One of the main arguments for its use is the fact that individuals who belong to the same social and geographic group share ecological expositions that make it highly likely that they have similar observed and unobserved characteristics². Mixed-effects models (MM) have been widely used in public health to model both geographical and non-geographical classifications of people. Some of the most studied classifications are neighborhood and social class, respectively. However, in recent years the applicability and pertinence of hierarchical MM to model geographic data have been discussed. A complete review of the main discussions can be found in a recent paper by Owen et al.³.

In the context of the geographical classification of people, hierarchical MM captures vertical dependence (hierarchy) but omits horizontal dependence (proximity) between geographic units⁴. In other words, these models do not take into consideration the spatial configuration of geographic units. The validity of interpretations about geographical phenomena that arise from non-spatial mixed-effects analyses have been questioned⁵. Even when the standard hierarchical MM only takes into account individual membership to a single geographic unit, interpretations tend to be done in terms of geographical effects. An alternative to standard MM are multiple membership models, which recognize that an individual can

belong to multiple groups simultaneously. In a geographical setting, multiple membership is supposed to model the effect of surrounding geographic units on an individual membership to a single geographic unit. Nevertheless, membership models, both single and multiple, are closer to the concept of place, which recognizes group membership, but may not operationalize space well, a concept that incorporates the interactions between places⁶.

Spatial econometricians have also developed techniques to answer similar questions about geographical effects, while public health researchers have mainly opted for multilevel models. One of the most widely used alternatives in spatial econometrics is the simultaneous autoregressive model (SAR). SAR models⁷ and other similar ones recognize the spatial arrangement of geographic units but do not allow for the identification of local differences. They only permit the evaluation of the global model. In addition, SAR models exclude the possibility of evaluating interactions between the characteristics of individuals and geographic units, which may lead to ecological fallacies.

These two methodological approaches respond to two different questions that could become complementary. The hierarchical MM model's main focus is on vertical dependency and vertical effects, while the spatial econometric model's main focus is on horizontal dependency and horizontal effects. Dong⁸ identified how the failure to simultaneously consider both effects in research settings, where hierarchical classification has a spatial nature, may lead to confounding. He found that a hierarchical MM could erroneously treat a horizontal interaction effect as a vertical contextual effect; and a spatial econometrics model could erroneously treat a vertical contextual effect as a horizontal interaction effect, which also may lead to confounding. The alternative, to simultaneously consider both effects, are recent. Main developments can be found in spatial econometrics with spatially varying autoregressive models (SVSAR)⁹, and in public health with spatial random effects models (SREMM)¹⁰. SREMM are basically mixed-effects models with spatial consideration defined within a conditional autoregressive prior (CAR). However, those models tend to be applied to aggregate data and have disease mapping purposes rather than analytical purposes with individual data. An exception can be found in recent work by Dong et al. where the model is extended to individual nested data⁵.

In this paper, a hybrid approach to geographic disparity analyses with elements from social epidemiology and spatial econometrics were tested in an attempt to face spatial relation misspecification and ecological bias, the main limitations of each individual methodology. In the context of social and geographic disparities, an appropriate measurement of the magnitude of the between-group variability in mortality, and the identification of factors that explain it, are highly relevant. They can provide a deeper understanding of the determinants of mortality disparities and consequently, help encourage appropriate public health and policy decisions, which can focus on areas for intervention and the prioritization of resources.

Thus, the aim of this paper is to analyze the conceptual and technical differences between three definitions of spatial relations within the Bayesian mixed-effects frameworks: classical multilevel definition, spatial multiple membership definition and conditional autoregressive definition, with an illustration of the estimation of geographic disparities in early neonatal mortality in Colombia from 2011 to 2014.

METHODS

DESIGN

A registry based cross-sectional study was conducted. All births recorded in the Vital Statistics registry for the period 2011 – 2014 with complete information on newborn sex, mother's municipality of residence, and residential area type were considered as part of the study population.

VARIABLES AND DATA SOURCE

Data on births and early neonatal deaths (within first seven days of life) that occurred during the study period were obtained from the Vital Statistics registry¹¹. 2,666,483 births were recorded, 5,073 (0.2%) birth certificates had quality problems and were excluded. A total of 2,661,410 births were analyzed. 13,478 early neonatal deaths were registered, 930 (7.0%) death certificates had quality problems and were excluded. A total of 12,538 deaths were included. Births and deaths were classified by the mother's municipality of residence according to Colombian political-administrative division in 1,122 geographic areas. Residential area type and a proxy of Gross Domestic Product – GDP — were considered as individual and area-level variables, respectively.

DEFINITIONS OF SPATIAL RELATION BETWEEN MUNICIPALITIES

Three Bayesian mixed-effects logistic models of newborns (y_i) cross-classified in municipalities (π_i) were constructed, one for each definition of spatial relations. Definitions are presented in cross-classified models notation¹² to facilitate their comparison. For all models, individual-level early neonatal death, the dependent variable y_i , was assumed to be binomially distributed with a probability equal to π_i (Equation 1), and regression equations were defined with a random intercept (β_{0i}) and fixed slopes ($\beta_{k,k \neq 0}$) for individual and area-level variables, as presented in Equation 2.

$$y_i \sim \text{Binomial}(1, \pi_i) \quad (1)$$

$$\text{logit}(\pi_i) = \beta_{0i} + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} \quad (2)$$

Municipal early neonatal mortality rates were estimated using local intercepts (β_{0i}). The area-level residual for each municipality ($u_{0,m(i)}^{(2)}$) was assumed to be normally distributed with the mean and variance specified according to each definition of spatial relations, as presented in Figure 1. The main differences between the definitions are related to assumptions about spatial

autocorrelation (or independence) of area-level residuals and the way to handle them. Only these differences are highlighted for each definition. Technical details can be consulted in the references provided.

CLASSICAL MULTILEVEL MODEL (CM)

The CM model¹³ implies a hierarchical structure with no explicit definition for spatial relations, ignoring the geographic arrangement of municipalities. In this model the estimation of residuals (Figure 1, Equation 4) only takes into consideration the municipality where each newborn resides, and independence between area-level residuals is assumed in the distribution function (Figure 1, Equation 5).

SPATIAL MULTIPLE MEMBERSHIP MODEL (SMM)

Similar to the CM model, the SMM model¹⁴ assumes independence between area-level residuals in the distribution function, thus it uses the same parameters that the CM model does (Figure 1, Equation 5). However, the spatial dimension is operationalized by weighting the area-level residuals of neighboring municipalities when estimating each local intercept (Figure 1, Equation 3). Consequently, SMM indirectly relates municipalities through the

Model	Local intercept	Parameters for the distribution of area-level residuals
SMM	$\beta_{0i} = \beta_0 + \sum_{j \in Nm(i)} w_j^{(2)} u_{0,m(i)}^{(2)} + e_{0i}$ [3]	$u_{0,m(i)}^{(2)} \sim N(0, \Omega_u^{(2)})$ with $\Omega_u^{(2)} = \Omega_{u0,0}^{(2)}$ [5]
CM		
CAR	$\beta_{0i} = \beta_0 + u_{0,m(i)}^{(2)} + e_{0i}$ [4]	$u_{0,m(i)}^{(2)} \sim N\left(\bar{u}_{0,m(i)}^{(2)}, \frac{\Omega_u^{(2)}}{r_{m(i)}^{(2)}}\right)$ with $\Omega_u^{(2)} = \Omega_{u0,0}^{(2)}$, and [6] $\bar{u}_{0,m(i)}^{(2)} = \sum_{j \in Nm(i)} w_j^{(2)} u_{0j}^{(2)} / r_{m(i)}^{(2)}$

β_{0i} : local intercept for the municipality where the newborn i resides; β_0 : fixed intercept; $u_{0,m(i)}^{(2)}$: area-level residual for the municipality where the newborn i resides; e_{0i} : error term at the individual-level, $w_j^{(2)}$: area-level weight of the neighbor municipality j ; $j \in Nm(i)$: the neighborhood constructed for the municipality where the newborn i resides; $\Omega_u^{(2)}$: variance of the area-level residuals; $\Omega_{u0,0}^{(2)}$: sample variance of the area-level residuals; $\bar{u}_{0,m(i)}^{(2)}$: average of the area-level residuals with the conditional autoregressive prior constrain; $r_{m(i)}^{(2)}$: size of the neighborhood of the municipality where the newborn i resides. Equations are expressed with notation for cross-classified models¹².

Figure 1. Local intercept and parameter equations for the distribution of area-level residuals for each spatial relation definition.

cross-classification of newborns. Neighborhood was defined as all adjacent municipalities and each one was assigned a weight proportional to the size of the neighborhood.

SPATIAL EFFECTS WITH THE CONDITIONAL AUTOREGRESSIVE PRIOR MODEL (CAR)

The CAR model does not assume independence between area-level residuals like the CM and SMM models do. It models area-level spatial autocorrelation of residuals using a conditional autoregressive prior⁵ with functions for mean and variance, which involve the neighborhoods residuals (Figure 1, Equation 6). Similar to spatial autocorrelation that is modeled in the distribution function, the CAR model does not apply weights to area-level residuals like the SMM model does, but it uses the same equation for the CM model (Figure 1, Equation 4). The CAR model directly relates municipalities at the area-level. The same definition of neighborhood used for the SMM model was applied to the CAR model, but neighborhoods' weights were fixed to 1.

For the SMM and CAR models, the neighborhoods' matrix was constructed with the ArcGIS add-in Adjacency for WinBUGS¹⁵.

DATA ANALYSIS

Bayesian estimation specifications

Random models were estimated with second order linearization and empirical Bayes Penalized Quasi-Likelihood methods, as was recommended for modeling spatial data¹⁶. A weakly informative prior with Inverse-Gamma (0.001, 0.001) distribution was used for the variance hyperparameter and non-informative priors and normal distribution were used for fixed parameters based on maximum likelihood estimates. The posterior distributions of parameters were estimated through Markov Chain Monte Carlo Simulations with Gibbs sampling and Metropolis-Hastings updates. Chain convergence was assessed using visual inspection based on graphics of trace, kernel density and autocorrelation function. Additionally, effective sample size was expected to be > 200. After a burn-in period of 500 iterations, an additional 5,000 iterations chain was sufficient to fulfill the requirements for the CM and CAR models but not for the SMM.

The parameters for crude and adjusted spatial relations definition models with their 95% credible intervals (95%CrI) are presented and compared in terms of proportional reduction (crude parameter– adjusted parameter)/ crude parameter.

ESTIMATION OF GEOGRAPHIC DISPARITIES IN MORTALITY

Crude and adjusted early neonatal mortality rates (per 1,000 live births (LB)) were estimated from intercepts. Geographic disparities in mortality were estimated from the variance

component and are presented as Median Rate Ratio (MRR) and Interquartile Rate Ratio (IqRR). They are two epidemiological measures of heterogeneity proposed by social epidemiology^{17,18}.

ESTIMATION OF RATE RATIOS AND RATE DIFFERENCES FOR INDIVIDUAL AND AREA-LEVEL VARIABLES

Rate Ratios were estimated from regression coefficients. Rate Differences were estimated from predictions for median differences against the reference category, and are expressed as deaths per 1,000 LB.

MODEL FIT AND SPATIAL AUTOCORRELATION

Model fit was assessed with three fit statistics: Deviance Information Criterion (DIC)¹⁹, effective number of parameters²⁰ and effective sample size²¹. The adequacy of the definitions of spatial relations used to model spatial autocorrelation of area-level residuals was evaluated with Moran's I statistic. Three spatial patterns on residuals were defined as: clustered, over dispersed or not relevant.

Mixed-effects analyses were executed using MLwiN 2.34 (Bristol University, UK). Spatial autocorrelation analyses were executed using ArcMap 10.2.2 (ESRI, US).

ETHICAL CONSIDERATIONS

Approval for this study, protocol number 466/2015, was obtained from the research ethics committee of the Universidad CES, Medellin, Colombia.

RESULTS

Table 1 shows the estimates of geographic disparities with regard to mortality and rate ratios for individual and area-level variables for each spatial relations definition model.

ESTIMATION OF GEOGRAPHIC DISPARITIES IN MORTALITY

Crude early neonatal mortality calculated from raw data was 4.7 deaths per 1,000 LB. The classic multilevel model estimated crude early neonatal mortality as 5.1 deaths per 1,000 LB. Both SMM and CAR made different estimates of 5.2 and 4.7 deaths per 1,000 LB respectively. After adjusting for newborn sex, residential area type, and GDP, estimated mortality rates were 2.2, 4.0 and 1.2 per 1,000 LB according to the CM, SMM and CAR models. This implies proportional reductions of crude rates after an adjustment of 57%, 24% and

Table 1. Estimated rate ratios and epidemiological measures of inter-municipal heterogeneity (with 95% credible intervals) for early neonatal mortality according to each spatial relation definition, Colombia 2011-2014.

	Definition of spatial relation between municipalities					
	Classical Multilevel		Spatial Multiple Membership		Conditional Autoregressive	
	RR (95% CrI)		RR (95% CrI)		RR (95% CrI)	
	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted
Sex						
Female ^a		1.0		1.0		1.0
Male		1.3 (1.2 – 1.3)		1.3 (1.2 – 1.3)		1.3 (1.2 – 1.3)
Residential area						
Semi-urban ^a		1.0		1.0		1.0
Urban		1.5 (1.3 – 1.6)		1.4 (1.3 – 1.5)		1.5 (1.3 – 1.6)
Rural		1.4 (1.2 – 1.5)		1.3 (1.2 – 1.5)		1.4 (1.2 – 1.5)
Gross domestic product^b						
3,203.3 – 48,255.9 ^a		1.0		1.0		1.0
502.5 – 3203.1		1.0 (0.8 – 1.2)		0.7 (0.5 – 1.1)		1.2 (0.8 – 1.8)
199.3 – 502.2		1.2 (0.9 – 1.5)		0.6 (0.5 – 1.0)		1.3 (0.9 – 1.8)
132.2 – 198.9		1.1 (0.9 – 1.4)		0.7 (0.5 – 1.1)		1.3 (0.9 – 1.9)
71.6 – 131.9		1.4 (1.1 – 1.7)		0.7 (0.5 – 1.0)		1.4 (1.0 – 2.1)
30.6 – 71.3		1.6 (1.3 – 1.9)		0.9 (0.7 – 1.4)		1.6 (1.1 – 2.4)
0.3 – 30.1		1.9 (1.6 – 2.3)		1.0 (0.8 – 1.7)		1.9 (1.3 – 2.9)
Mortality rate ^c	5.1 (4.8 5.3)	2.2 (1.9 – 2.7)	5.2 (4.4 – 6.2)	4.0 (2.6 – 5.0)	4.7 (3.9 – 5.8)	1.2 (0.9 – 2.1)
Median RR	2.0 (1.9 2.1)	1.9 (1.8 – 2.0)	12.1 (9.9 – 14.6)	10.3 (8.5 – 12.6)	4.7 (4.2 – 5.3)	4.5 (4.1 – 5.0)
Interquartile RR	5.2 (4.7 5.7)	4.7 (4.2 – 5.2)	405.0 ^e	280.1 ^f	42.6 (32.6 – 55.2)	38.3 (29.5 – 49.2)
Effective sample	640	418	150	96	449	475
Effective parameters	660	651	537	529	691	690
Deviance Information	154889	154619	155054	154765	154879	154602
Spatial pattern of residuals ^d	Clustered	Clustered	Dispersed	Dispersed	Clustered	Not relevant

^aReference category; ^bUSD Millions; ^cPer thousand live births; ^dBased on Moran's I Statistic; ^e(95%CrI 250.4 – 638.8); ^f(95%CrI 172.8 – 446.1); RR: Rate Ratio; 95%CrI: 95% Credible interval.

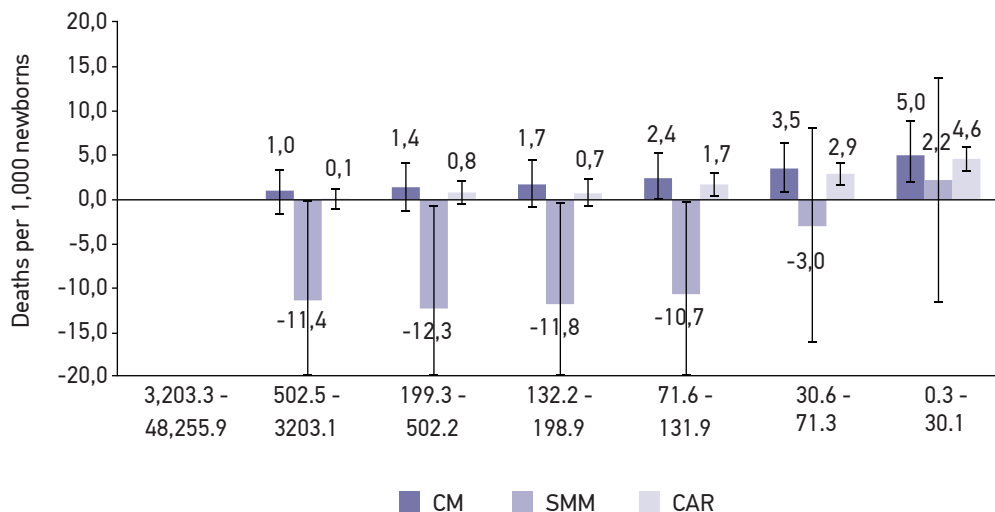
74% respectively. The CAR and SMM models also showed greater uncertainty in credible intervals than the CM model for both crude and adjusted rates.

According to the CAR model crude-MRR, relative difference in mortality is higher than 4.7 for half of the possible comparisons between municipalities. Crude-IqRR shows that for newborns residents of the top 25% highest mortality municipalities, mortality is 42.6 times higher than for newborns residents of the bottom 25% lowest mortality municipalities. Crude and adjusted MRR and IqRR were lower for CM than for CAR. Despite these differences between models, the proportional reduction of the geographic disparity measures after adjustment was similar: 4% for MRR and 10% for IqRR. The SMM model made substantially higher estimations for geographic disparity estimates and for proportional reduction after adjustment, 14% for MRR and 31% for IqRR, compared to both CM and CAR.

ESTIMATION OF RATE RATIOS AND RATE DIFFERENCES FOR INDIVIDUAL AND AREA-LEVEL VARIABLES

The estimates of the individual-level variable Rate Ratios were the same for all models. The estimated relative effect of GDP categories varied between models according to Rate Ratios. Both CM and CAR identified a social gradient, but classical multilevel median estimates were lower and credible intervals were narrower. SMM made an inconsistent estimate compared to CM and CAR.

Figure 2 shows the Rate Differences for municipal GDP categories (reference: 3,203.3 - 48,255.9 USD million). For the area-level variable, the CM estimates were higher than the



CM: Classical Multilevel; SMM: Spatial Multiple Membership; CAR: Conditional autoregressive.

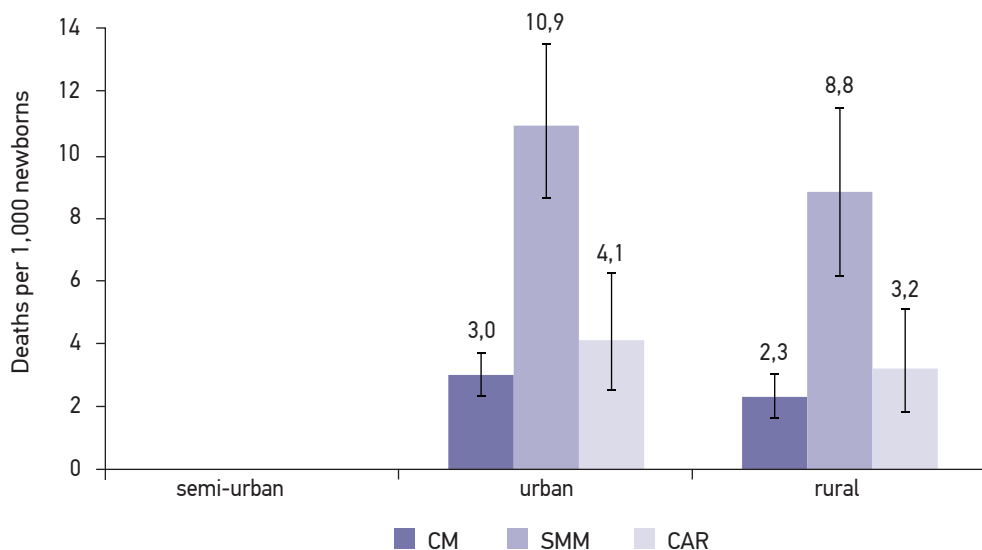
Figure 2. Estimated rate differences (with 95% credible intervals) between the area-level variable gross domestic product according to each spatial relation definition, Colombia 2011 – 2014.

CAR estimates, and they showed higher uncertainty. According to the median CAR estimates, municipalities with a GDP lower than USD 30.1 million have 4.6 more early neonatal deaths per 1,000 live births than municipalities with a higher GDP. In municipalities with GDP categories USD 30.6 – 71.3 and 71.6 – 131.9 millions, the median rate difference compared to reference category was 2.9 and 1.7 deaths per 1,000 live births. The SMM estimates are incoherent with CM and CAR, and show higher uncertainty.

Figure 3 shows the Rate Differences between newborn's residential area types (reference: semi-urban). For the individual-level variable, the CAR estimates are higher than the CM, around 1 death per 1,000 live births. The SMM made a considerably higher estimate of Rate Differences compared to both CAR and CM, and showed higher uncertainty.

MODEL FIT AND SPATIAL AUTOCORRELATION

In terms of model fit (Table 1), the CAR model showed the highest number of effective parameters and the lowest DIC for both crude and adjusted models. The CAR also showed the highest effective sample size for the adjusted models, but not between crude models where the CM model showed the highest value. Spatial autocorrelation of area-level residuals was properly modeled in the CAR adjusted model, but not in the CAR crude model. Crude and adjusted model patterns of area-level residuals were clustered for CM and dispersed for SMM.



CM: Classical Multilevel; SMM: Spatial Multiple Membership; CAR: Conditional autoregressive.

Figure 3. Estimated rate differences (with 95% credible intervals) between the individual-level variable residential area type according to each spatial relation definition, Colombia 2011-2014.

DISCUSSION

The greatest difference between the three definitions of spatial relations was observed for the area-level variance estimate, which is used to calculate MRR and IqRR—the epidemiological measures of geographic disparity. Differences can be explained by the spatial autocorrelation of area-level residuals. CM underestimation of geographic disparity may be explained by its lack of a process to model spatially autocorrelated variability. SMM overestimated the geographic disparity measures, because when it weighted the residuals (Figure 1, Equation 3), it produced dispersed residuals (beyond what was expected by chance). The CAR estimation of geographic disparity measures can be considered the most appropriate, since its prior for the distribution of residuals (Figure 1, Equation 6) produced spatially non-correlated residuals and showed the best model fit. Additionally, the CAR definition of spatial relations is a horizontal (proximity) operationalization of space²². The SMM definition of spatial relations is a vertical (hierarchy) operationalization that does not relate municipalities to each other, but rather individuals to multiple municipalities.

The similarity between the Rate Ratios estimated for sex and residential area type was expected as mixed-models estimates of individual-level fixed effects are intra-cluster specific, and therefore are not affected by differences in the estimation of area-level variance²³. This is not the case for Rate Differences estimated for individual-level variables, which were obtained by prediction, a process affected by area-level variance²⁴. A similar logic applies to area-level estimates of Rate Ratios and Rate Differences, as fixed effects are cluster specific²³. The differences observed between the CM and CAR models' GDP Rate Ratios may be explained by the weighting applied to the mean of the normal distribution assumed for the area-level residuals (Figure 1, Equation 6).

Uncertainty differences in all parameter estimates also must be cautiously interpreted. Even though, in every single case, the CM model produced narrower credible intervals than the CAR, this can be explained by CM underestimation of variance (as explained above). Ignoring autocorrelation leads to spurious certainty²⁵. For the CM model, autocorrelation should be considered when constructing credible intervals. One alternative for maximum likelihood estimation is the method for calculating confidence intervals for Rate Ratios between geographic units in the presence of spatial autocorrelation. This alternative was proposed by Zhu²⁶.

It is important to highlight that even when the CM model showed lower model fit than the CAR, and produced spatially auto-correlated residuals, this model cannot be discarded. Two considerations should be made:

1. as said earlier in this paper, the conceptual distinction between place and space must be carefully thought out⁶. Classical multilevel can be a more adequate model in some research scenarios where group membership is of primary interest rather than spatial distribution;
2. CAR implies the possibility of adjacent units, which is more common in secondary data analyses as censuses or vital statistics registries. Other data sources such as health,

demographic surveys, or even primary data from complex sampling designs may not be adequate because they do not necessarily have adjacent territories. In those cases, the CM model still can be used, but conclusions must be confined to place interpretations and interpretation of spatial effects should be avoided.

Further research from an epidemiological point of view is needed for a deeper understanding of those effects in order to clarify the differential mechanism associated with space and place as geographic dimensions of health. The main hypothesis proposed from these study results is that space and place, as geographic dimensions of health, might behave in a similar fashion as age and cohort do as time dimensions of health, where age-effects may be overlaid by cohort-effects²⁷ leading to confounding. Space-effects may be overlaid by place-effects but more research is needed to test this hypothesis and to provide a structural definition for bias. Further research from the point of view of biostatistics is also required. It is known²⁸ that for datasets in which low variances are possible, inverse-gamma prior yields an improper posterior distribution for the limit of $\epsilon \rightarrow 0$, making inferences very sensitive to the values of ϵ . This is partly due to the complex dependence of the variance marginal likelihood on the distribution of data between area-level units. The CAR model has a more complex variance²² than CM and specific research on proper priors for it is necessary.

The public health applications of spatial effects models are worth mentioning. Several papers using the definition of spatial relations with the CAR model can be found for disease mapping and ecological analysis²⁹⁻³², but their applications and comparisons for geographic disparities research with data at an individual level are scarce³³. The main strength of this paper is that the definitions of spatial relations are performed in a common general framework which makes their comparison more straightforward. Also, differences in terms of epidemiological measures, estimation, and potential confounding, rather than technical issues, are the main focus.

This study has a number of limitations:

1. The inclusion of the CAR prior was not enough to properly model the spatial nature in the crude model. This is relevant for disease and mortality mapping, since it suggests that even when a proper spatial model is used to represent crude rates distribution with a map, spatial-nonstationarity may not be sufficiently modeled. According to findings that were not presented, even when adjusting mortality rates only for sex using the CAR model, the area-level residuals showed no spatial autocorrelation. However, improved model specification is needed in order to gain a greater understanding of the spatial non-stationarity processes underlying spatial distribution of mortality rates³⁴;
2. It is important to note that the poor performance of the spatial multiple membership model could be due to the distribution of weights among the neighborhood. As such, the impact of weighting schemes in the SMM model fit must be explored in depth. In this regard, the CAR model has the advantage of an unambiguous definition of weights, which is equal to one for adjacent municipalities, and to zero otherwise⁵. Additionally, the SMM chain convergence for 5,000 iterations was not optimal;

3. Short Markov Chains were used because of the high computational demands of the Bayesian mixed effect models with the CAR and SMM structures, and few independent variables were considered. However, as the main purpose was to compare different definitions of spatial relation in terms of conceptual and technical aspects, this basic model was enough for illustrative purposes and had enough fit evidence and chain stability according to ESS, spatial autocorrelation of residuals, and visual inspection of chain convergence. The persistent spatial autocorrelation of the residuals in CM could be diminished by including more covariates, but it is important to highlight that in spatial data modeling the spatial non-stationarity must not be considered as model misfit or noise, but rather as a main process of relevance that should not be adjusted, but rather analytically considered⁴.

CONCLUSION

The definition of spatial relations between municipalities based on the conditional autoregressive prior showed the best performance according to both fit statistics and spatial pattern analyses of area-level residuals. The use of the CAR model in the Bayesian mixed-effects framework, allows for the possibility to simultaneously:

1. estimate individual-level probability of death;
2. estimate mortality rates for geographic units;
3. estimate geographic disparity measures;
4. adjust those estimates for both individual and area-level variables, while properly considering spatial arrangement of the geographic units.

This approach diminishes ecological bias and avoids spatial misspecification in geographic disparities analyses. If further research increases the validity arguments for this approach, it may be an important contribution to epidemiological design as an alternative to ecological studies in the study of geographic disparities in mortality.

REFERENCES

1. Goldstein H. *Multilevel Statistical Models*. 4th ed. New York: Wiley; 2010.
2. Browne W, Goldstein H. MCMC sampling for a multilevel model with non independent residuals within and between cluster units. *J Educ Behav Stat*. 2010;35(4):453-73.
3. Owen G, Harris R, Jones K. Under examination multilevel models, geography and health research. *Prog Hum Geogr*. 2015;40(3):394-412.
4. Dong G, Harris R. Spatial autoregressive models for geographically hierarchical data structures. *Geogr Anal*. 2015;47(2):173-91.
5. Dong G, Ma J, Harris R, Pryce G. Spatial random slope multilevel modeling using multivariate conditional autoregressive models: A case study of subjective travel satisfaction in Beijing. *Ann Am Assoc Geogr*. 2016;106(1):19-35.
6. Arcaya M, Brewster M, Zigler CM, Subramanian SV. Area variations in health: A spatial multilevel modeling approach. *Health Place*. 2012;18(4):824-31.
7. Pinault L, Crouse D, Jerrett M, Brauer M, Tjepkema M. Spatial associations between socioeconomic groups and NO₂ air pollution exposure within three large Canadian cities. *Environ Res*. 2016;147:373-82.

8. Dong G. Simultaneous modelling of contextual and spatial interaction effects: towards a synthesis of spatial and multilevel modelling [Doctoral thesis]. [Bristol, U.K.]: University of Bristol; 2014.
9. Mukherjee C, Kasibhatla PS, West M. Spatially varying SAR models and Bayesian inference for high-resolution lattice data. *Ann Inst Stat Math*. 2014;66(3):473-94.
10. Congdon P. A spatially adaptive conditional autoregressive prior for area health data. *Stat Methodol*. 2008;5(6):552-63.
11. National Administrative Department of Statistics (CO). Vital statistics database [Internet]. 2012 [cited on 2016 Apr. 7]. Available from: <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/nacimientos-y-defunciones>
12. Browne WJ, Goldstein H, Rasbash J. Multiple membership multiple classification (MMMC) models. *Statistical Modelling*. 2001;1(2):103-24.
13. Duncan C, Jones K, Moon G. Context, composition and heterogeneity: Using multilevel models in health research. *Soc Sci Med* 1998;46(1):97-117.
14. Chandola T, Clarke P, Wiggins RD, Bartley M. Who you live with and where you live: Setting the context for health using multiple membership multilevel models. *J Epidemiol Community Health*. 2005;59(2):170-5.
15. Upper Midwest Environmental Sciences Center. Adjacency for WinBUGS [Internet]. 2012. [cited 2016 Mar. 14]. Available from: http://www.umesc.usgs.gov/management/dss/adjacency_tool.html
16. MacNab YC, Lin Y. On empirical Bayes penalized quasi-likelihood inference in GLMMs and in Bayesian disease mapping and ecological modeling. *Comput Stat Data Anal*. 2009;53(8):2950-67.
17. Merlo J, Chaix B, Ohlsson H, Beckman A, Johnell K, Hjerpe P, et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: Using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *J Epidemiol Community Health*. 2006;60(4):290-7.
18. Larsen K, Merlo J. Appropriate assessment of neighborhood effects on individual health: integrating random and fixed effects in multilevel logistic regression. *Am J Epidemiol*. 2005;161(1):81-8.
19. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Ser B Stat Methodol*. 2002;64(4):583-639.
20. Leeuw J de, Meijer E, Eds. *Handbook of multilevel analysis*. New York: Springer; 2008.
21. Morita S, Thall PF, Müller P. Determining the effective sample size of a parametric prior. *Biometrics*. 2008;64(2):595-602.
22. Wall MM. A close look at the spatial structure implied by the CAR and SAR models. *J Stat Plan Inference*. 2004;121(2):311-24.
23. Snijders TAB, Bosker RJ. *Multilevel analysis: an introduction to basic and advanced multilevel modeling*. 2^a ed. Los Angeles: Sage; 2012.
24. Rasbash J, Charlton C, Jones K, Pillingier R. *Manual supplement for MLwiN*. Version 2.36 [Internet]. 2016 [cited on 2016 Mar. 14]. Available from: <http://www.bristol.ac.uk/cmm/media/software/mlwin/downloads/manuals/2-36/supplement-web.pdf>
25. Biggs R, Carpenter SR, Brock WA. Spurious Certainty: How Ignoring Measurement Error and Environmental Heterogeneity May Contribute to Environmental Controversies. *BioScience*. 2009;59(1):65-76.
26. Zhu L, Pickle LW, Pearson JB. Confidence intervals for rate ratios between geographic units. *Int J Health Geogr*. 2016;15(1):44.
27. Schilling OK. Cohort- and age-related decline in elder's life satisfaction: is there really a paradox? *Eur J Ageing*. 2005;2(4):254-63.
28. Gelman A. Prior distributions for variance parameters in hierarchical models (Comment on Article by Browne and Draper). *Bayesian Analysis*. 2006;1:515-34.
29. Kim H, Sun D, Tsutakawa RK. A bivariate Bayes method for improving the estimates of mortality rates with a twofold conditional autoregressive model. *J Am Statistical Assoc*. 2001;96(456):1506-21.
30. Neelon B, Anthopolos R, Miranda ML. A spatial bivariate probit model for correlated binary data with application to adverse birth outcomes. *Stat Methods Med Res*. 2014;23(2):119-33.
31. Porter AT, Holan SH, Wikle CK. Bayesian semiparametric hierarchical empirical likelihood spatial models. *J Statist Plan Inference*. 2015;165:78-90.
32. Lee D. A comparison of conditional autoregressive models used in Bayesian disease mapping. *Spat Spatio-Temporal Epidemiol*. 2011;2(2):79-89.
33. Dasgupta P, Cramb SM, Aitken JF, Turrell G, Baade PD. Comparing multilevel and Bayesian spatial random effects survival models to assess geographical inequalities in colorectal cancer survival: a case study. *Int J Health Geogr*. 2014;13:36.
34. Fuglstad G-A. *Modelling Spatial non-stationarity* [Doctoral thesis]. [Trondheim (NO)]: Norwegian University of Science and Technology; 2014.

Received on: 06/05/2016

Final version presented on: 11/01/2017

Accepted on: 17/02/2017