Biometrics 64, 1–??
December 2008

DOI: 10.1111/j.1541-0420.2005.00454.x

Supporting Information for "Simultaneous spatial smoothing and outlier detection using penalized regression, with application to childhood obesity surveillance from electronic health records" by Choi, Hanrahan, Norton, and Zhao

Young-Geun Choi[1], Lawrence P. Hanrahan[2], Derek Norton[3], and Ying-Qi Zhao[4,*],

[1]Department of Statistics, Sookmyung Women's University,

Seoul, South Korea.

[2]Department of Family Medicine and Community Health, University of Wisconsin-Madison,

Madison, Wisconsin, U.S.A.

[3]Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison,

Madison, Wisconsin, U.S.A.

[4]Public Health Sciences Division, Fred Hutchinson Cancer Research Center,

Seattle, Washington, U.S.A.

**email:* yqzhao@fredhutch.org

Received October 2007. *Revised February* 2008. *Accepted March* 2008.

Appendix A. Proof of Proposition 1

From the construction of the $\boldsymbol{\alpha}$- and $\boldsymbol{\gamma}$-steps, it is obvious that $\phi(\boldsymbol{\alpha}^{(t)}, \boldsymbol{\beta}^{(t)}, \boldsymbol{\gamma}^{(t)}) \geqslant \phi(\boldsymbol{\alpha}^{(t+1)}, \boldsymbol{\beta}^{(t)}, \boldsymbol{\gamma}^{(t)})$ and $\phi(\boldsymbol{\alpha}^{(t+1)}, \boldsymbol{\beta}^{(t+1)}, \boldsymbol{\gamma}^{(t)}) \geqslant \phi(\boldsymbol{\alpha}^{(t+1)}, \boldsymbol{\beta}^{(t+1)}, \boldsymbol{\gamma}^{(t+1)})$. To guarantee

$$\phi(\boldsymbol{\alpha}^{(t+1)}, \boldsymbol{\beta}^{(t)}, \boldsymbol{\gamma}^{(t)}) \geqslant \phi(\boldsymbol{\alpha}^{(t+1)}, \boldsymbol{\beta}^{(t+1)}, \boldsymbol{\gamma}^{(t)}),$$

we verify a more general statement. Lemma 1 indicates that if an objective function $\psi(\cdot)$ consists of a smooth convex loss function $l(\cdot)$ plus a convex (and possibly non-differentiable) penalty $P(\cdot)$, one can descend ψ by minimizing a surrogate function of ψ in which the loss part l is replaced by the local quadratic approximation of l.

LEMMA 1: *Let $l(\boldsymbol{t})$ be a twicely-differentiable convex function and $P(\boldsymbol{t})$ be a convex function that is not necessarily differentiable. Define $\psi(\boldsymbol{t}) := l(\boldsymbol{t}) + P(\boldsymbol{t})$. Let $\boldsymbol{t}^0$ be a fixed value in the domain of ψ. Define*

$$\tilde{l}(\boldsymbol{t};\boldsymbol{t}^0) := l(\boldsymbol{t}^0) + \nabla_{\boldsymbol{t}} l(\boldsymbol{t}^0)^T(\boldsymbol{t}-\boldsymbol{t}^0) + \frac{1}{2}(\boldsymbol{t}-\boldsymbol{t}^0)^T \nabla^2_{\boldsymbol{tt}} l(\boldsymbol{t}^0)(\boldsymbol{t}-\boldsymbol{t}^0), \qquad \tilde{\psi}(\boldsymbol{t};\boldsymbol{t}^0) := \tilde{l}(\boldsymbol{t};\boldsymbol{t}^0) + P(\boldsymbol{t}),$$

and $\boldsymbol{t}^* := \operatorname{argmin}_{\boldsymbol{t}} \tilde{\psi}(\boldsymbol{t};\boldsymbol{t}^0)$. Consider

$$\boldsymbol{t}^1 := h^*\boldsymbol{t}^* + (1-h^*)\boldsymbol{t}^0, \quad \textit{where } h^* := \operatorname*{argmin}_{h\in[0,1]} \psi\left(h\boldsymbol{t}^* + (1-h)\boldsymbol{t}^0\right).$$

If $\boldsymbol{t}^0$ is not a minimizer of $\psi(\boldsymbol{t})$ (i.e., $\psi(\boldsymbol{t}^0) < \psi(\boldsymbol{t}^*)$), then such h^* exists and $\psi(\boldsymbol{t}^1) < \psi(\boldsymbol{t}^0)$.

The proof is essentially the same as the proof of Proposition 1 in Lee et al. (2016). They assumed $P(\boldsymbol{t}) = \lambda\|\boldsymbol{t}\|_1$ for some λ, which can be easily extended to a general convex function $P(\cdot)$.

Proof. Note that $\psi(\boldsymbol{t})$ and $\tilde{\psi}(\boldsymbol{t};\boldsymbol{t}^0)$ have the same penalty function, and $\nabla_{\boldsymbol{t}} l(\boldsymbol{t}) = \nabla_{\boldsymbol{t}} \tilde{l}(\boldsymbol{t}^0;\boldsymbol{t}^0)$. In addition, $\psi(\boldsymbol{t})$ and $\tilde{\psi}(\boldsymbol{t};\boldsymbol{t}^0)$ have the same Karush-Kuhn-Tucker first-order optimality conditions at $\boldsymbol{t} = \boldsymbol{t}^0$. Then the assumption $\boldsymbol{t}^0 \neq \operatorname{argmin}_{\boldsymbol{t}} \psi(\boldsymbol{t})$ is equivalent to $\boldsymbol{t}^0 \neq \operatorname{argmin}_{\boldsymbol{t}} \tilde{\psi}(\boldsymbol{t};\boldsymbol{t}^0)$.

Consequently, $\tilde{\psi}(\boldsymbol{t}^*;\boldsymbol{t}^0) < \tilde{\psi}(\boldsymbol{t}^0;\boldsymbol{t}^0)$. Now let $h \in (0,1]$ and $\boldsymbol{t}^h = h\boldsymbol{t}^* + (1-h)\boldsymbol{t}^0$. The convexity of $\tilde{\psi}(\cdot;\boldsymbol{t}^0)$ implies

$$\tilde{\psi}(\boldsymbol{t}^h;\boldsymbol{t}^0) \leqslant h\tilde{\psi}(\boldsymbol{t}^*;\boldsymbol{t}^0) + (1-h)\tilde{\psi}(\boldsymbol{t}^0;\boldsymbol{t}^0),$$

which yields

$$\frac{\tilde{\psi}(\boldsymbol{t}^h;\boldsymbol{t}^0) - \tilde{\psi}(\boldsymbol{t}^0;\boldsymbol{t}^0)}{h} \leqslant \tilde{\psi}(\boldsymbol{t}^*;\boldsymbol{t}^0) - \tilde{\psi}(\boldsymbol{t}^0;\boldsymbol{t}^0) < 0.$$

Furthermore,

$$\begin{aligned} & \frac{\psi(\boldsymbol{t}^h) - \psi(\boldsymbol{t}^0)}{h} \\ = \; & \frac{\tilde{\psi}(\boldsymbol{t}^h;\boldsymbol{t}^0) - \tilde{\psi}(\boldsymbol{t}^0;\boldsymbol{t}^0)}{h} - \frac{\tilde{l}(\boldsymbol{t}^h;\boldsymbol{t}^0) - \tilde{l}(\boldsymbol{t}^0;\boldsymbol{t}^0)}{h} + \frac{l(\boldsymbol{t}^h) - l(\boldsymbol{t}^0)}{h} \\ \leqslant \; & \tilde{\psi}(\boldsymbol{t}^*;\boldsymbol{t}^0) - \tilde{\psi}(\boldsymbol{t}^0;\boldsymbol{t}^0) - \frac{\tilde{l}(\boldsymbol{t}^h;\boldsymbol{t}^0) - \tilde{l}(\boldsymbol{t}^0;\boldsymbol{t}^0)}{h} + \frac{l(\boldsymbol{t}^h) - l(\boldsymbol{t}^0)}{h}. \end{aligned}$$

As $h \to 0^+$, $\tilde{\psi}(\boldsymbol{t}^*;\boldsymbol{t}^0) - \tilde{\psi}(\boldsymbol{t}^0;\boldsymbol{t}^0) - \{\tilde{l}(\boldsymbol{t}^h;\boldsymbol{t}^0) - \tilde{l}(\boldsymbol{t}^0;\boldsymbol{t}^0)\}/h$ converges to $-\nabla_{\boldsymbol{t}}\tilde{l}(\boldsymbol{t}^0;\boldsymbol{t}^0) + \nabla_{\boldsymbol{t}} l(\boldsymbol{t})$, which vanish to zero by the construction of $\tilde{l}$. Therefore, we have

$$\limsup_{h\to 0^+} \frac{\psi(\boldsymbol{t}^h) - \psi(\boldsymbol{t}^0)}{h} \leqslant \tilde{\psi}(\boldsymbol{t}^*;\boldsymbol{t}^0) - \tilde{\psi}(\boldsymbol{t}^0;\boldsymbol{t}^0) < 0.$$

Therefore, there exists at least one $h \in (0,1]$ such that $\psi(\boldsymbol{t}^h) < \psi(\boldsymbol{t}^0)$. $\psi(\boldsymbol{t}^1) < \psi(\boldsymbol{t}^0)$ subsequently according to the construction of $\boldsymbol{t}^1$.

Appendix B. Modified objective function adjusted for weights

Denote the final weight as w_{ij} for (ij)-th subject. The modified objective function ϕ is defined by

$$\phi^w(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}) = -\text{loglik}^w(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}) + P_{\lambda_1}(\boldsymbol{\beta}) + Q^w_{\lambda_2}(\boldsymbol{\gamma}), \tag{A.1}$$

where

$$\begin{aligned}\text{loglik}^w(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}) &= \frac{1}{w_{\cdot\cdot}}\sum_{i=1}^{K}\sum_{j=1}^{n_i} w_{ij} I(R_{ij}=1)\cdot\Big[\log\{1+\exp(\boldsymbol{Z}_{ij}^T\boldsymbol{\alpha}_1+\boldsymbol{X}_i^T\boldsymbol{\alpha}_2 \\ &\quad +\beta_i+\gamma_i)\} - Y_{ij}(\boldsymbol{Z}_{ij}^T\boldsymbol{\alpha}_1+\boldsymbol{X}_i^T\boldsymbol{\alpha}_2+\beta_i+\gamma_i)\Big],\end{aligned}$$

and

$$Q_{\lambda_2}^w(\boldsymbol{\gamma}) = \frac{1}{w_{\cdot\cdot}}\sum_{i=1}^{K} w_{i\cdot} q_{\lambda_2}(\gamma_i),$$

with $w_{i\cdot} = \sum_j w_{ij} I(R_{ij}=1)$ and $w_{\cdot\cdot} = \sum_i \sum_j w_{ij} I(R_{ij}=1)$. The case of complete data can be understood as $w_{ij}=1$ and $R_{ij}=1$ for all i and j. Algorithm S1 in Appendix C describes the alternating minimization algorithm adjusted for the weight. The result of Proposition 1 remains the same as long as the dataset contains at least one obese and non-obese subject observed for all locations. The modified BIC reflecting the weight is

$$\text{BIC}^{w*}(\lambda_1,\lambda_2) = -2w_{\cdot\cdot}\cdot\text{loglik}^w(\widehat{\boldsymbol{\alpha}},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\gamma}}) + \text{DF}\cdot(\log w_{\cdot\cdot}+1),$$

where $\text{loglik}^w(\widehat{\boldsymbol{\alpha}},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\gamma}})$ is given in (A.1) and DF is given in (6) in Section 3.4.

Appendix C. Optimization algorithm

Algorithm S1 describes the proposed alternating minimization algorithm solving (A.1).

Algorithm S1 An alternating minimization algorithm for (A.1)

---

require: Arrays $\{Y_{ij}\}$, $\{\boldsymbol{Q}_{ij}\}$, $\{R_{ij}\}$ and $\{w_{ij}\}$, $i = 1, \ldots, n_i$, $j = 1, \ldots, K$, array $\{\rho_{i_1,i_2}\}$, scalar λ_1 and scalar λ_2, tolerance level $\epsilon = 10^{-6}$

initialize $\boldsymbol{\alpha}^{(0)}$, $\boldsymbol{\beta}^{(0)}$, $\boldsymbol{\gamma}^{(0)}$, $\phi^{(0)} = \phi(\boldsymbol{\alpha}^{(0)}, \boldsymbol{\beta}^{(0)}, \boldsymbol{\gamma}^{(0)})$

define $w_{i\cdot} \leftarrow \sum_j w_{ij} I(R_{ij} = 1)$ and $w_{\cdot\cdot} \leftarrow \sum_i \sum_j w_{ij} I(R_{ij} = 1)$

while $\frac{|\phi^{(t+1)} - \phi^{(t)}|}{\max\{1, |\phi^{(t)}|\}} > \epsilon$ **do**

(1. Updating $\boldsymbol{\alpha}$)

1-1. $\mu_{ij}^{(t)} \leftarrow \beta_i^{(t)} + \gamma_i^{(t)}$ for $j = 1, \ldots, n_i$, $i = 1, \ldots, K$.

1-2. Run a logistic regression, without intercept, for $N = \sum_{i=1}^K n_i$ individuals with response $\{Y_{ij}\}$, predictor $\{\boldsymbol{Q}_{ij}\}$, offset $\{\mu_{ij}^{(t)}\}$ and weight $\{w_{ij}\}$.

1-3. Assign the results from Steps 1-1/1-2 to $\boldsymbol{\alpha}_i^{(t+1)}$.

(2. Updating $\boldsymbol{\beta}$)

2-1. $\theta_{ij}^{(t)} \leftarrow \boldsymbol{Q}_{ij}^T \boldsymbol{\alpha}^{(t+1)} + \gamma_i^{(t)}$ for $j = 1, \ldots, n_i$, $i = 1, \ldots, K$.

2-2. $a_i^{(t)} \leftarrow \sum_{j=1}^{n_i} w_{ij} \left[\frac{\exp\left(\beta_i^{(t)} + \theta_{ij}^{(t)}\right)}{\left\{1 + \exp\left(\beta_i^{(t)} + \theta_{ij}^{(t)}\right)\right\}^2} \right]$ for $i = 1, \ldots, K$.

2-3. $b_i^{(t)} \leftarrow \beta_i^{(t)} - \frac{1}{a_i^{(t)}} \sum_{j=1}^{n_i} w_{ij} \left[\frac{\exp\left(\beta_i^{(t)} + \theta_{ij}^{(t)}\right)}{1 + \exp\left(\beta_i^{(t)} + \theta_{ij}^{(t)}\right)} - Y_{ij} \right]$ for $i = 1, \ldots, K$.

2-4. Solve

$$\tilde{\boldsymbol{\beta}} \leftarrow \underset{\boldsymbol{\beta} \in \mathbb{R}^K}{\operatorname{argmin}} \left[\frac{1}{2w_{\cdot\cdot}} \sum_{i=1}^K a_i^{(t)} \left(\beta_i - b_i^{(t)} \right)^2 + \lambda_1 \sum_{i_1 < i_2} \rho_{i_1,i_2} |\beta_{i_1} - \beta_{i_2}| \right].$$

2-5. If $\phi(\boldsymbol{\alpha}^{(t+1)}, \tilde{\boldsymbol{\beta}}, \boldsymbol{\gamma}^{(t)}) \leqslant \phi(\boldsymbol{\alpha}^{(t+1)}, \boldsymbol{\beta}^{(t)}, \boldsymbol{\gamma}^{(t)})$, then $\boldsymbol{\beta}^{(t+1)} \leftarrow \tilde{\boldsymbol{\beta}}$. Otherwise, $\boldsymbol{\beta}^{(t+1)} \leftarrow \tilde{h}\tilde{\boldsymbol{\beta}} + (1 - \tilde{h})\boldsymbol{\beta}^{(t)}$, where

$$\tilde{h} = \underset{h \in [0,1]}{\operatorname{argmin}} \, \phi \left(\boldsymbol{\alpha}^{(t+1)}, h\tilde{\boldsymbol{\beta}} + (1 - h)\boldsymbol{\beta}^{(t)}, \boldsymbol{\gamma}^{(t)} \right).$$

(3. Updating $\boldsymbol{\gamma}$)

3-1. $\nu_{ij}^{(t)} \leftarrow \boldsymbol{Q}_{ij}^T \boldsymbol{\alpha}^{(t+1)} + \beta_i^{(t+1)}$ for $j = 1, \ldots, n_i$, $i = 1, \ldots, K$.

3-2. For $i = 1, \ldots, K$:

$$\gamma_i^{(t+1)} \leftarrow \underset{\gamma}{\operatorname{argmin}} \left[\sum_{j=1}^{n_i} w_{ij} \left[\log \left\{ 1 + \exp\left(\gamma + \nu_{ij}^{(t)} \right) \right\} - Y_{ij} \left(\gamma + \nu_{ij}^{(t)} \right) \right] + w_{i\cdot} q_{\lambda_2}(\gamma) \right].$$

4. $\phi^{(t+1)} \leftarrow \phi(\boldsymbol{\alpha}^{(t+1)}, \boldsymbol{\beta}^{(t+1)}, \boldsymbol{\gamma}^{(t+1)})$

end while

return $(\boldsymbol{\alpha}^{(t+1)}, \boldsymbol{\beta}^{(t+1)}, \boldsymbol{\gamma}^{(t+1)})$

Appendix D. Additional simulation results

In Section D.1, we display additional plots for the simulations considered in Section 4 in the main body. In Section D.2, we report the performance of methods when the outcome Y could be missing.

D.1. Additional plots for missing outcome

For a detailed comparison of the performances in outlier detection, we additionally investigated true positive rate (TPR or sensitivity) and true negative rate (TNR or specificity) of the considered methods as in Figures S1 and S2. We recall that the TPR and TNR are defined by

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \qquad \mathrm{TNR} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}},$$

where TP, TN, FP and FN were defined in Section 4. The figures suggests that our methods consistently outperforms other methods in TPR while the TNRs of all the methods are comparable.

[Figure 1 about here.]

[Figure 2 about here.]

D.2. Additional setting for missing outcome

We considered three missing mechanisms of the outcome Y: (M1). $\mathbb{P}(R_{ij} = 1|Z_1, Z_2)$ was set to 0.7; (M2). $\mathbb{P}(R_{ij} = 1|Z_1, Z_2) = \text{logit}^{-1}(2 + 2Z_1 + 2Z_2)$; or (M3). $\mathbb{P}(R_{ij} = 1|Z_1, Z_2) = \text{logit}^{-1}(2+4Z_1Z_2)$. M1 stands for missing completely at random, while M2 and M3 represents missing at random. We estimate the missing probability by the logistic regression with predictor Z_1 and Z_2. As anticipated, there are more biases in the estimated coefficients in the presence of missingness, but the proposed method outperformed the competitors overall. Table S1, Figures S3, S4, S5 and S6 summarize the simulation results.

[Table 1 about here.]

[Figure 3 about here.]

[Figure 4 about here.]

[Figure 5 about here.]

[Figure 6 about here.]

Appendix E. Additional plot for the real data analysis

Figure S7 displays a choropleth map of the area-level raw childhood obesity prevalence rates.

[Figure 7 about here.]

References

Lee, S., Kwon, S., and Kim, Y. (2016). A modified local quadratic approximation algorithm for penalized optimization problems. *Computational Statistics and Data Analysis*, 94:275–286.

Figure S1. TPR, varying the number of outliers over 1000 replications.

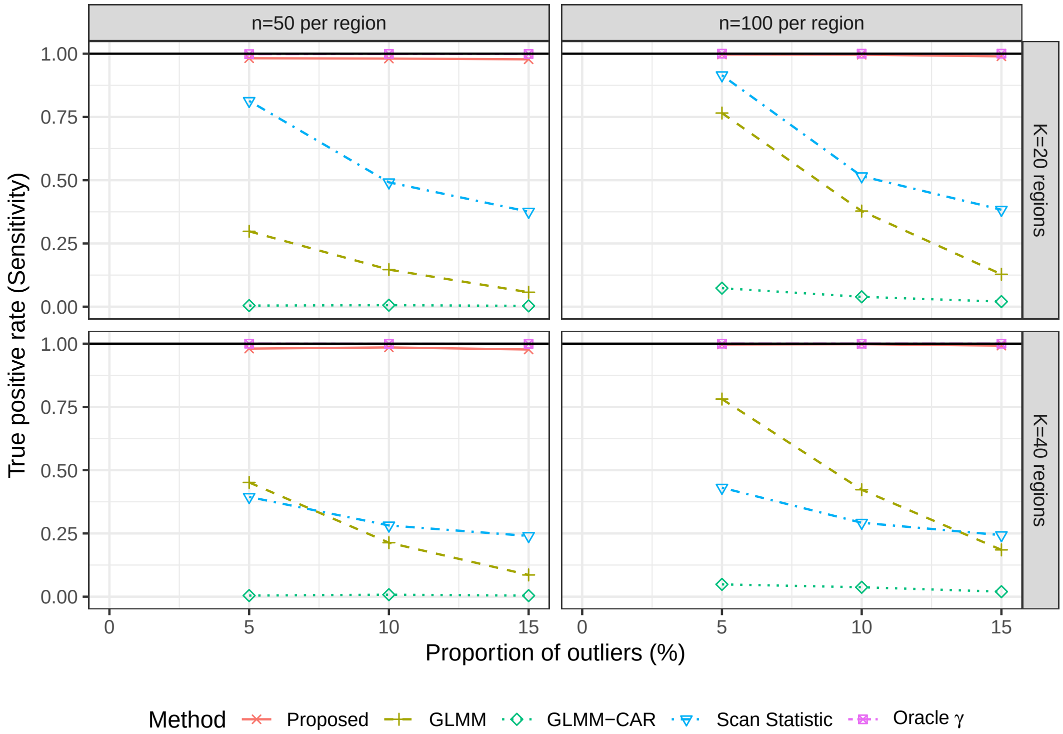

Figure S2. TNR, varying the number of outliers over 1000 replications.

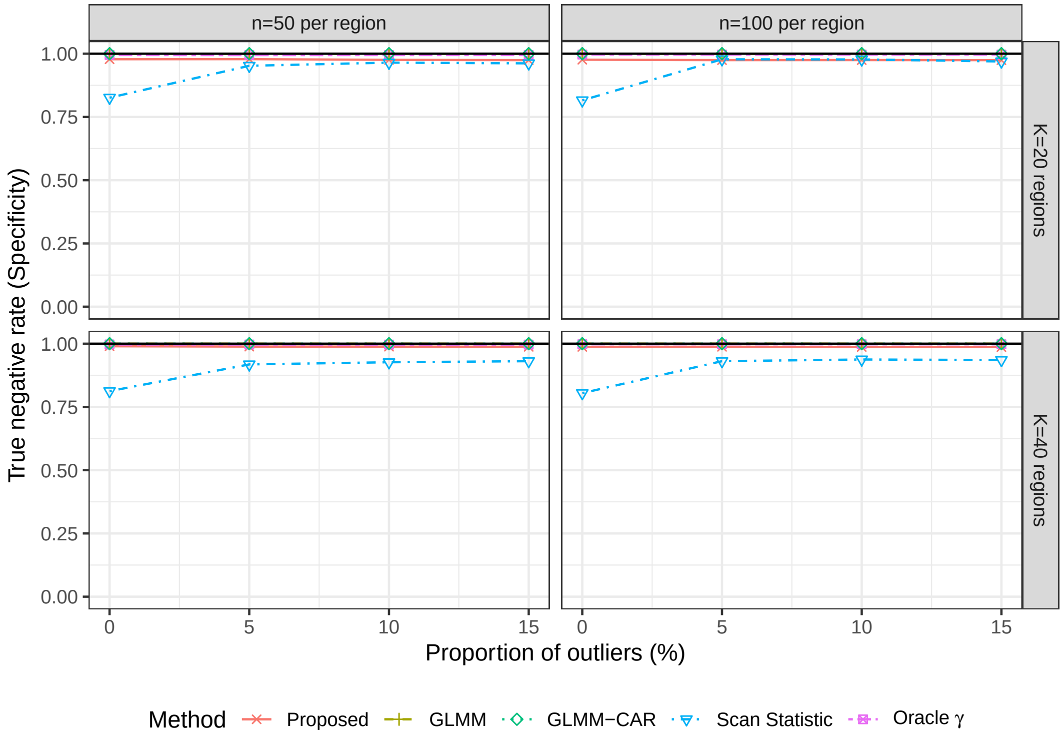

Figure S3. MCC, varying Y-missing mechanisms over 1000 replications.

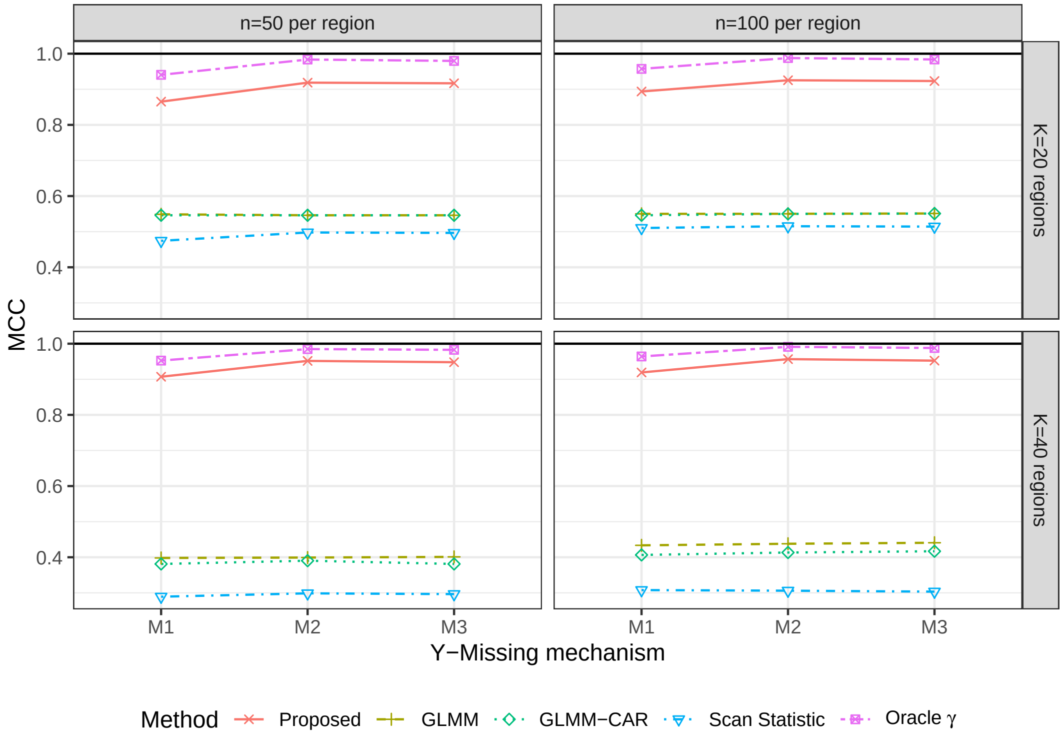

Figure S4. TPR, varying Y-missing mechanisms over 1000 replications.

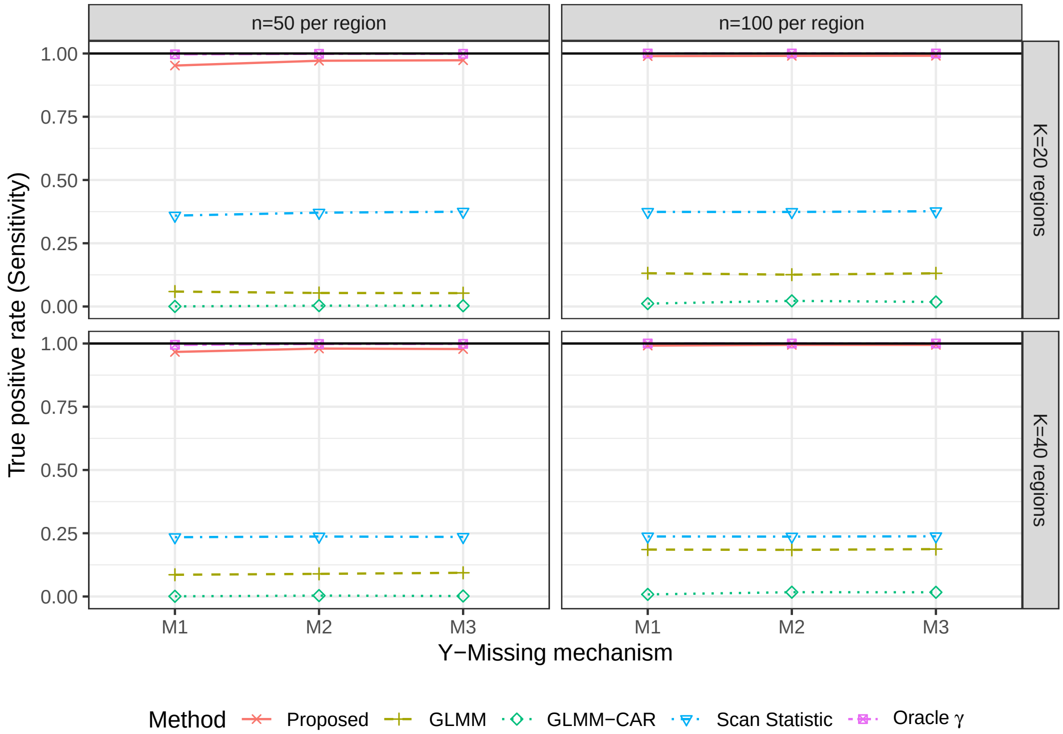

Figure S5. TNR, varying Y-missing mechanisms over 1000 replications.

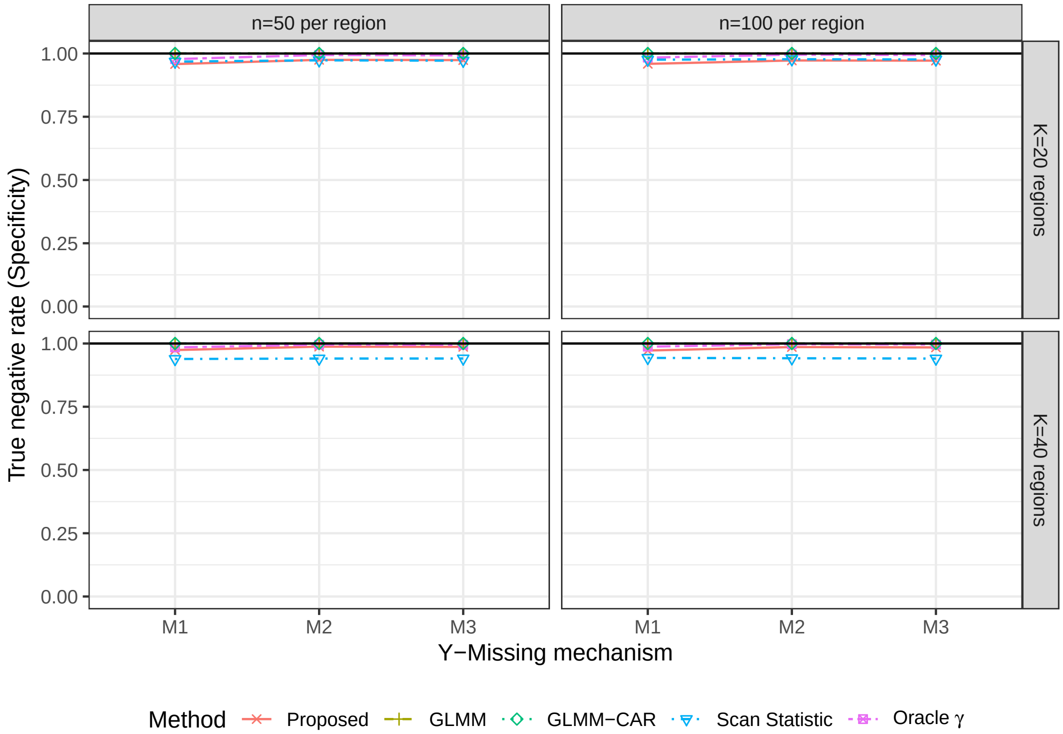

Figure S6. RMSE of $\widehat{\boldsymbol{\beta}}$, varying Y-missing mechanisms over 1000 replications.

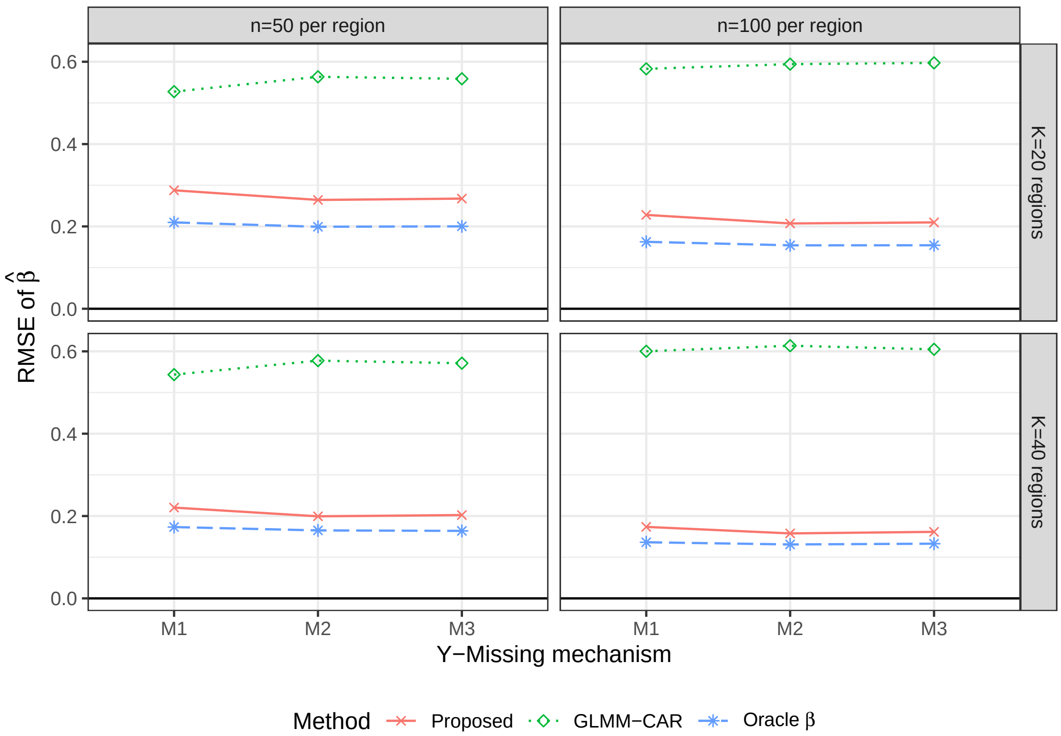

Figure S7. A choropleth map of the area-level raw childhood obesity prevalence rates.

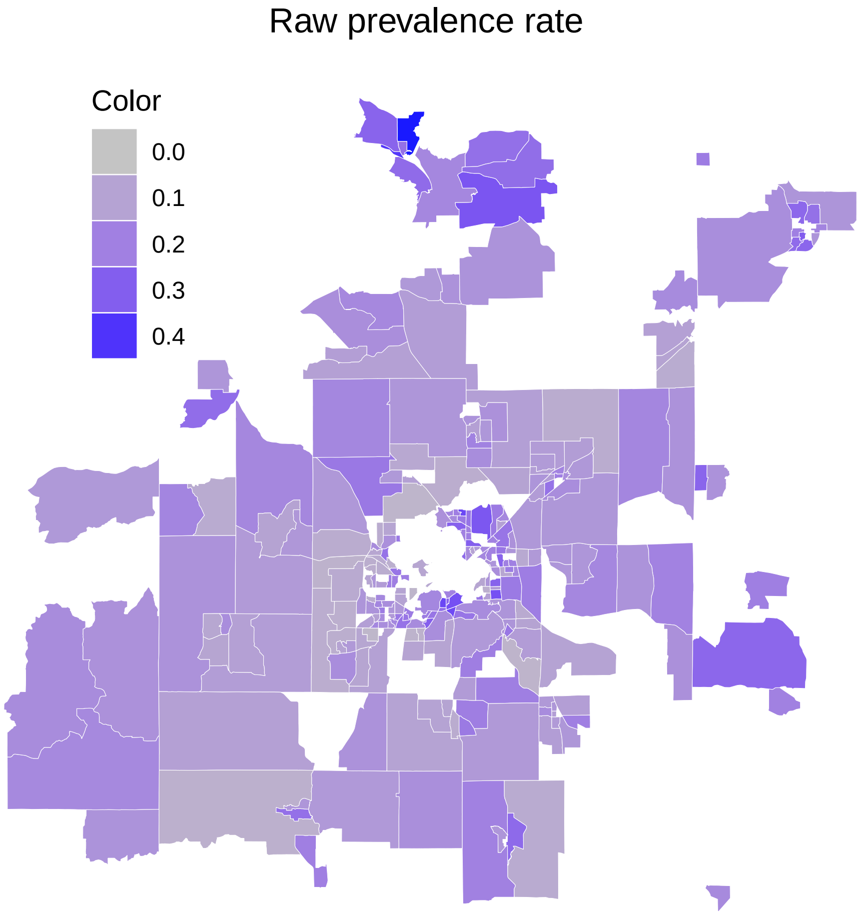

Table S1

Biases (± standard errors) and bootstrap coverage probabilities of $\widehat{\alpha}_1$ and $\widehat{\alpha}_2$, and RMSE of $\{\widehat{p}_i\}_{i=1}^K$, varying Y-missing mechanisms over 1000 replications.

Y-missing model	Method	$n = 50$ per region					$n = 100$ per region				
		$\widehat{\alpha}_1$ Bias	CP	$\widehat{\alpha}_2$ Bias	CP	RMSE of $\{\widehat{p}_i\}$	$\widehat{\alpha}_1$ Bias	CP	$\widehat{\alpha}_2$ Bias	CP	RMSE of $\{\widehat{p}_i\}$
$K = 20$ regions											
M1	Proposed	-.007 ± .011	.943	-.005 ± .017	.958	.066	.002 ± .007	.960	-.017 ± .012	.960	.047
	GLMM	-.009 ± .011	.933	-.039 ± .024	.617	.075	.001 ± .007	.958	-.025 ± .024	.458	.055
	GLMM-CAR	-.008 ± .011	.930	-.038 ± .024	.634	.075	.001 ± .007	.957	-.028 ± .024	.471	.055
	Oracle $\boldsymbol{\alpha}$	.001 ± .009	.931	.000 ± .009	.937	.019	.002 ± .006	.948	-.004 ± .006	.961	.013
M2	Proposed	-.006 ± .009	.937	-.011 ± .014	.941	.058	.003 ± .006	.957	-.022 ± .010	.922	.041
	GLMM	-.007 ± .009	.941	-.037 ± .023	.534	.065	.002 ± .006	.960	-.027 ± .023	.415	.047
	GLMM-CAR	-.008 ± .009	.939	-.035 ± .024	.546	.065	.002 ± .006	.960	-.027 ± .024	.424	.047
	Oracle $\boldsymbol{\alpha}$	-.000 ± .007	.945	.003 ± .007	.950	.016	.004 ± .005	.955	-.005 ± .005	.938	.011
M3	Proposed	-.004 ± .009	.924	-.010 ± .014	.945	.058	.002 ± .006	.953	-.019 ± .011	.920	.042
	GLMM	-.006 ± .009	.929	-.038 ± .023	.565	.066	.001 ± .006	.952	-.028 ± .023	.433	.048
	GLMM-CAR	-.007 ± .009	.926	-.036 ± .024	.563	.066	.001 ± .006	.953	-.029 ± .024	.419	.048
	Oracle $\boldsymbol{\alpha}$	.001 ± .008	.927	.002 ± .007	.954	.016	.003 ± .005	.953	-.004 ± .005	.952	.011
$K = 40$ regions											
M1	Proposed	-.008 ± .007	.964	.006 ± .010	.983	.057	.003 ± .005	.945	.005 ± .007	.975	.041
	GLMM	-.009 ± .007	.968	.013 ± .017	.622	.074	.002 ± .005	.945	.012 ± .017	.449	.055
	GLMM-CAR	-.008 ± .007	.971	.008 ± .017	.624	.074	.002 ± .005	.942	.009 ± .017	.453	.055
	Oracle $\boldsymbol{\alpha}$	-.005 ± .006	.948	.005 ± .006	.950	.014	.002 ± .004	.952	-.000 ± .004	.935	.010
M2	Proposed	-.005 ± .006	.954	.001 ± .008	.969	.047	-.001 ± .004	.950	.003 ± .006	.952	.035
	GLMM	-.005 ± .006	.960	.010 ± .016	.523	.064	-.002 ± .004	.948	.012 ± .017	.402	.047
	GLMM-CAR	-.005 ± .006	.960	.007 ± .017	.529	.064	-.002 ± .004	.950	.010 ± .017	.409	.047
	Oracle $\boldsymbol{\alpha}$	-.002 ± .005	.943	.003 ± .005	.947	.012	-.001 ± .003	.953	.002 ± .003	.957	.008
M3	Proposed	-.005 ± .006	.950	.000 ± .009	.981	.048	.000 ± .004	.958	.003 ± .006	.960	.036
	GLMM	-.006 ± .006	.952	.011 ± .016	.533	.066	-.001 ± .004	.956	.008 ± .017	.408	.048
	GLMM-CAR	-.006 ± .006	.956	.008 ± .017	.539	.066	-.001 ± .004	.960	.005 ± .017	.421	.048
	Oracle $\boldsymbol{\alpha}$	-.003 ± .005	.935	.003 ± .005	.941	.012	.000 ± .003	.958	.001 ± .004	.958	.008