Online Supplementary Materials for "Bayesian analysis of time-series data under case-crossover designs: posterior equivalence and inference" by Shi Li, Bhramar Mukherjee, Stuart Batterman and Malay Ghosh

Web Appendix A: Equivalence results between time-series analysis and case-crossover design

A.1: Frequentist equivalence using full likelihood

From (5), $\log(L^T_{full}(\boldsymbol{\beta}, \nu)) = \sum_{t=1}^{T} Y_t \Big[\nu_t + \boldsymbol{\beta}^\top \boldsymbol{X}_t - \sum_{s\in W(t)} \log\{1+\exp(\nu_t + \boldsymbol{\beta}^\top \boldsymbol{X}_s)\}\Big]$ and then the estimating equation for $\boldsymbol{\beta}$ is

$$\begin{aligned}
U^T_{full}(\boldsymbol{\beta}) &= \sum_{t=1}^{T} Y_t \boldsymbol{X}_t - \sum_{t=1}^{T}\sum_{s=1}^{T} Y_t \frac{I(s\in W(t))\exp(\nu_t + \boldsymbol{\beta}^\top \boldsymbol{X}_s)\boldsymbol{X}_s}{1+\exp(\nu_t + \boldsymbol{\beta}^\top \boldsymbol{X}_s)} \\
&= \sum_{t=1}^{T} Y_t \boldsymbol{X}_t - \sum_{t=1}^{T}\sum_{s=1}^{T} Y_s \frac{I(t\in W(s))\exp(\nu_s + \boldsymbol{\beta}^\top \boldsymbol{X}_t)\boldsymbol{X}_t}{1+\exp(\nu_s + \boldsymbol{\beta}^\top \boldsymbol{X}_t)} \\
&= \sum_{t=1}^{T} Y_t \boldsymbol{X}_t - \sum_{t=1}^{T} \boldsymbol{X}_t \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t)\Big\{\sum_{s=1}^{T} \frac{Y_s I(s\in R(t))\exp(\nu_s)}{1+\exp(\nu_s + \boldsymbol{\beta}^\top \boldsymbol{X}_t)}\Big\} \\
&= \sum_{t=1}^{T} \boldsymbol{X}_t \Big\{Y_t - \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t) \sum_{s\in R(t)} \frac{Y_s \exp(\nu_s)}{1+\exp(\nu_s + \boldsymbol{\beta}^\top \boldsymbol{X}_t)}\Big\},
\end{aligned}$$

where $R(t)$ is the set of days that contain day t in their reference window. For SBD and TSD but not more generally, $R(t) = W(t)$ (Lu and Zeger 2007). Comparing with the log-linear model estimating equation corresponding to β derived from (7)

$$U_{ll}(\boldsymbol{\beta}) = \sum_{t=1}^{T} \boldsymbol{X}_t \Big\{Y_t - \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t + S_t)\Big\},$$

So, if $\widehat{S}_t(\nu, \boldsymbol{\beta}) = \log(\sum_{s\in R(t)} Y_s \exp(\nu_s)/\{1+\exp(\nu_s + \boldsymbol{\beta}^\top \boldsymbol{X}_t)\})$, then $U^T_{full}(\boldsymbol{\beta})$ will provide the same estimate for $\boldsymbol{\beta}$ as $U_{ll}(\boldsymbol{\beta})$. Under a TSD, while the conditional likelihood approach or an equivalent log-linear model would only allow the risk changes abruptly among different time stratifications, the full likelihood approach does not require such constraint because $\widehat{S}_{t'}(\boldsymbol{\nu}, \boldsymbol{\beta})$ is not necessarily equal to $\widehat{S}_t(\boldsymbol{\nu}, \boldsymbol{\beta})$ for $t' \in W(t)$ and $t' \neq t$.

A.2: Bayesian equivalence using conditional likelihood

Proof of Theorem 1: From (7),

$$L_{ll}(\boldsymbol{\beta}, S_t) \;\propto\; \prod_{k=1}^{K} \prod_{t:\, t\in ts(k)} \big\{ \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t + S'_k)\big\}^{Y_t} \exp\{-\exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t + S'_k)\}.$$

Let $\varphi_k = \exp(S'_k)$. The marginal posterior distribution of $\boldsymbol{\beta}$ derived from $L_{ll}(\boldsymbol{\beta}, S_t)$ is

$$\begin{aligned}
&\pi(\boldsymbol{\beta} \mid \boldsymbol{X}, \boldsymbol{Y}) \propto \int \pi(\boldsymbol{\beta})\pi(S'_1, ..., S'_K) L_{ll}(\boldsymbol{\beta}, S_t) dS'_1 \cdots dS'_K \\
\propto\;& \pi(\boldsymbol{\beta}) \int \prod_{k=1}^{K} \prod_{t:\, t\in ts(k)} \big\{ \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t + S'_k)\big\}^{Y_t} \exp\{-\exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t + S'_k)\} dS'_1 \cdots dS'_K \\
=\;& \pi(\boldsymbol{\beta}) \big[\prod_{t=1}^{T} \big\{ \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t)\big\}^{Y_t} \big] \prod_{k=1}^{K} \int \varphi_k^{-1} \prod_{t:\, t\in ts(k)} \big[\varphi_k^{Y_t} \exp\{-\varphi_k \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t)\}\big] d\varphi_k \\
=\;& \pi(\boldsymbol{\beta}) \big[\prod_{t=1}^{T} \big\{ \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t)\big\}^{Y_t} \big] \prod_{k=1}^{K} \int \big[\varphi_k^{\sum_{t:\, t\in ts(k)} Y_t - 1} \exp\{-\varphi_k \sum_{t:\, t\in ts(k)} \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t)\}\big] d\varphi_k \\
=\;& \pi(\boldsymbol{\beta}) \big[\prod_{t=1}^{T} \big\{ \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t)\big\}^{Y_t} \big] \prod_{k=1}^{K} \{ \sum_{t:\, t\in ts(k)} \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t)\}^{-\sum_{t\in ts(k)} Y_t} \\
=\;& \pi(\boldsymbol{\beta}) \big[\prod_{t=1}^{T} \big\{ \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t)\big\}^{Y_t} \big] \prod_{k=1}^{K} \prod_{t:\, t\in ts(k)} \{ \sum_{s\in W(t)} \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_s)\}^{-Y_t} \\
=\;& \pi(\boldsymbol{\beta}) \big[\prod_{t=1}^{T} \big\{ \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t)\big\}^{Y_t} \big] \prod_{t=1}^{T} \{ \sum_{s\in W(t)} \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_s)\}^{-Y_t} \\
=\;& \pi(\boldsymbol{\beta}) \prod_{t=1}^{T} \Big\{ \frac{\exp(\boldsymbol{\beta}^\top \boldsymbol{X}_t)}{\sum_{s\in W(t)} \exp(\boldsymbol{\beta}^\top \boldsymbol{X}_s)} \Big\}^{Y_t} \\
=\;& \pi(\boldsymbol{\beta}) L_{cc}(\boldsymbol{\beta}),
\end{aligned}$$

which is the marginal posterior distribution of $\boldsymbol{\beta}$ derived from $L_{cc}(\boldsymbol{\beta})$.

A.3: Bayesian equivalence using full likelihood

Let $y_{s\cdot t} = y_{s1t} + y_{s0t}$, and let $\Phi_{st} = \exp(\phi_{st})$. Then

$$
\begin{aligned}
& \pi(\boldsymbol{\nu}, \boldsymbol{\beta} \mid \boldsymbol{X}, \boldsymbol{Y}) \propto \int \pi(\boldsymbol{\phi}, \boldsymbol{\nu}, \boldsymbol{\beta}) L_p(\boldsymbol{\phi}, \boldsymbol{\nu}, \boldsymbol{\beta}) d\boldsymbol{\phi} \propto \pi(\boldsymbol{\nu}, \boldsymbol{\beta}) \prod_{t=1}^{T} \big\{ \exp(\nu_t + \boldsymbol{\beta}^\top \boldsymbol{X}_t) \big\}^{Y_t} \\
& \times \int \pi(\boldsymbol{\phi}) \prod_{s=1}^{T} \prod_{t=1}^{T} [\exp(\phi_{st})]^{y_{s\cdot t}} \exp\{-\exp(\phi_{st})[1 + \exp(\nu_s + \boldsymbol{\beta}^\top \boldsymbol{X}_t)]\} d\boldsymbol{\phi} \\
= \; & \pi(\boldsymbol{\nu}, \boldsymbol{\beta}) \prod_{t=1}^{T} \big\{ \exp(\nu_t + \boldsymbol{\beta}^\top \boldsymbol{X}_t) \big\}^{Y_t} \prod_{s=1}^{T} \prod_{t=1}^{T} \int \Phi_{st}^{y_{s\cdot t}-1} \exp\{-\Phi_{st}[1 + \exp(\nu_s + \boldsymbol{\beta}^\top \boldsymbol{X}_t)]\} d\Phi_{st} \\
= \; & \pi(\boldsymbol{\nu}, \boldsymbol{\beta}) \prod_{t=1}^{T} \big\{ \exp(\nu_t + \boldsymbol{\beta}^\top \boldsymbol{X}_t) \big\}^{Y_t} \prod_{s=1}^{T} \prod_{t=1}^{T} \Big\{ \frac{1}{1 + \exp(\nu_s + \boldsymbol{\beta}^\top \boldsymbol{X}_t)} \Big\}^{y_{s\cdot t}} \\
= \; & \pi(\boldsymbol{\nu}, \boldsymbol{\beta}) \prod_{t=1}^{T} \big\{ \exp(\nu_t + \boldsymbol{\beta}^\top \boldsymbol{X}_t) \big\}^{Y_t} \prod_{s=1}^{T} \prod_{t=1}^{T} \Big\{ \frac{1}{1 + \exp(\nu_s + \boldsymbol{\beta}^\top \boldsymbol{X}_t)} \Big\}^{y_s I(s \in R(t))} \\
= \; & \pi(\boldsymbol{\nu}, \boldsymbol{\beta}) \prod_{t=1}^{T} \big\{ \exp(\nu_t + \boldsymbol{\beta}^\top \boldsymbol{X}_t) \big\}^{Y_t} \prod_{t=1}^{T} \prod_{s=1}^{T} \Big\{ \frac{1}{1 + \exp(\nu_t + \boldsymbol{\beta}^\top \boldsymbol{X}_s)} \Big\}^{y_t I(s \in W(t))} \\
= \; & \pi(\boldsymbol{\nu}, \boldsymbol{\beta}) \prod_{t=1}^{T} \Big[\big\{ \exp(\nu_t + \boldsymbol{\beta}^\top \boldsymbol{X}_t) \big\}^{Y_t} \prod_{s \in W(t)} \Big\{ \frac{1}{1 + \exp(\nu_t + \boldsymbol{\beta}^\top \boldsymbol{X}_s)} \Big\}^{y_t} \Big] \\
= \; & \pi(\boldsymbol{\nu}, \boldsymbol{\beta}) \prod_{t=1}^{T} \Big[\frac{\exp(\nu_t + \boldsymbol{\beta}^\top \boldsymbol{X}_t)}{\prod_{s \in W(t)} \{1 + \exp(\nu_t + \boldsymbol{\beta}^\top \boldsymbol{X}_s)\}} \Big]^{Y_t} \\
= \; & \pi(\boldsymbol{\nu}, \boldsymbol{\beta}) L_{full}^{T}(\boldsymbol{\nu}, \boldsymbol{\beta})
\end{aligned}
$$

Web Appendix B: Computational details for Bayesian inference

B.1: Metropolis-Hastings within Gibbs algorithm

Sampling of $\boldsymbol{\beta}$ under conditional likelihood formulation: We take conditional likelihood (3) under the shared exposure as an example. For mutually independent normal priors $\boldsymbol{\beta} \sim N(\boldsymbol{\mu}_\beta, \sigma_\beta^2 \boldsymbol{I}_p)$, the joint posterior distribution of $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)^\top$ is not a standard distribution. Let $\pi(\boldsymbol{\theta} \mid \cdot)$ denote the full conditional distribution as a function of $\boldsymbol{\theta}$ given the data and all other parameters. The posterior distribution $\pi(\boldsymbol{\beta} \mid \boldsymbol{X}, \boldsymbol{Y})$ can be obtained using a Gibbs sampler through the following full conditional distributions,

$$\pi(\beta_r|\cdot) = \exp\Big\{-\frac{(\beta_r-\mu_{\beta_r})^2}{2\sigma^2_{\beta_r}}\Big\}\prod_{t=1}^{T}\Big\{\frac{\exp(\beta_r X_{tr})}{\sum_{s\in W(t)}\exp(\boldsymbol{\beta}^\top \boldsymbol{X}_s)}\Big\}^{Y_t}, \quad r=1,...,p.$$

To sample from $\pi(\beta_r \mid \cdot)$, we followed these steps of a Metropolis-Hastings algorithm:

- Step 1. Start with initial value $\boldsymbol{\beta}^{(0)} = (\beta_1^{(0)},...,\beta_p^{(0)})^\top$.
- Step 2. For $r = 1,...,p$, at the k-th iteration with the current value as $\boldsymbol{\beta}^{(k)} = (\beta_1^{(k)},...,\beta_{r-1}^{(k)}, \beta_r^{(k-1)},...,\beta_p^{(k-1)})^\top$. Generate a new value β_r^* from a candidate density $g(\beta_r)$ and replace $\beta_r^{(k)}$ by β_r^* with probability $\min\Big\{1, \frac{\pi(\beta_r^*|\cdot)g(\beta_r^{(k-1)})}{\pi(\beta_r^{(k-1)}|\cdot)g(\beta_r^*)}\Big\}$. We chose the candidate density $g(\beta_r)$ as the prior density $\pi(\beta_r)$. Since the full conditional density $\pi(\beta_r \mid \cdot) \propto \pi(\beta_r)L(\beta_r \mid \cdot)$, the acceptance probability reduces to $\min\Big\{1, \frac{L(\beta_r^*|\cdot)}{L(\beta_r^{(k-1)}|\cdot)}\Big\}$. $\beta_r^{(k)}, 1,...,p$, is updated accordingly.
- Step 3. Run the chain with 10,000 iterations.

Sampling of $\boldsymbol{\nu}$ and $\boldsymbol{\beta}$ under full likelihood formulation: We take full likelihood model $L^T_{full}(\boldsymbol{\beta},\boldsymbol{\nu})$ (5) under the shared exposure as an example. To sample from the posterior distribution of $\boldsymbol{\nu}$ and $\boldsymbol{\beta}$, we adopted a componentwise Metropolis-Hastings algorithm. The full conditional distributions used are:

$$\pi(\beta_r|\cdot) \propto \exp\Big\{-\frac{(\beta_r-\mu_{\beta_r})^2}{2\sigma^2_{\beta_r}}\Big\}\prod_{t=1}^{T}\Big\{\frac{\exp(\beta_r X_{tr})}{\prod_{s\in W(t)}[1+\exp(\nu_t+\boldsymbol{\beta}^\top \boldsymbol{X}_s)]}\Big\}^{y_t}, \quad r=1,....,p.$$

$$\pi(\nu_t|\cdot) \propto \Big\{\frac{1}{\prod_{s\in W(t)}[1+\exp(\nu_t+\boldsymbol{\beta}^\top \boldsymbol{X}_s)]}\Big\}^{y_t} \times$$

$$\Big\{\frac{\alpha}{T-1+\alpha}\frac{\exp(\mu y_t + \frac{y_t^2\sigma^2}{2})}{\sqrt{2\pi\sigma^2}} \times \exp\Big[-\frac{(\nu_t-\mu-y_t\sigma^2)^2}{2\sigma^2}\Big] + \frac{\exp(y_t\nu_t)}{T-1+\alpha}\sum_{s=1,\, s\neq t}^{T} I(\nu_s=\nu_t)\Big\},$$

At each iteration, we first update the value of $\boldsymbol{\beta}$ similarly as described above, and then move on to the cycle for $\boldsymbol{\nu}$ with the updated value of $\boldsymbol{\beta}$ substituted. Particularly, given current values of $\boldsymbol{\beta}$, $\boldsymbol{\nu}$ is updated in the following way:

- Step 1. As a metropolis-Hastings step, ν_t^* was drawn from the candidate distribution of $\pi(\nu_t \mid \boldsymbol{\nu}_{-t})$, namely from $\frac{\alpha}{T-1+\alpha}N(\mu,\sigma^2) + \frac{1}{T-1+\alpha}\sum_{s=1,\, s\neq t}^{T} I(\nu_s=\nu_t)$. In particular, one either get a distinct value for ν_t^* from the normal component with probability $\frac{\alpha}{T-1+\alpha}$ or get a draw of ν_t^* with equal probability from the current set of the $(T-1)$ entries of $\boldsymbol{\nu}_{-t}$. We adopt the algorithm in to generate observations from this candidate density.

- Step 2. Set the new value of ν_t to ν_t^*, with acceptance probability min$\left\{1, \frac{L(\nu_t^*|\cdot)}{L(\nu_t^{'}|\cdot)}\right\}$, where $\nu_t^{'}$ is the current working value of ν_t.

- Step 3. Repeat steps 1-2 5 times and consider the last of these updates of ν_t as the value of $\nu_t^{(k)}$, say at the k-th iteration.

- Step 4. Repeat steps 1-3 for all $\nu_t^{(k)}$ for $t = 1, ..., T$. One complete iteration of the Markov chain consists of the foregoing updates for both the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\nu}$. Given current values of $\boldsymbol{\beta}^{(k)}$ and $\boldsymbol{\nu}^{(k)}$, we can go to the next iteration for $\boldsymbol{\beta}^{(k+1)}$ and $\boldsymbol{\nu}^{(k+1)}$. We run the chain with 10,000 iterations.

B.2: Prior choices under the simulation study

Assume the informative prior on β has the form $\beta \sim N(\mu_\beta, \sigma_\beta^2)$. According to the ad-hoc prior eliciting strategy described in section 6 for the DAMAT study, when $\beta^* = 0.1$ we a priori postulated a 95% confidence interval (1.02, 1.15) for $\exp(\beta)$, and solved for the approximated values of $(\mu_\beta, \sigma_\beta)$ as $(0.05, 0.02)$. Thus our informative prior was chosen as $N(0.08, 0.03^2)$ when $\beta^* = 0.1$. Similarly, when $\beta^* = 1$ we presumed a 95% confidence interval $(0.4, 1.2)$ for $\exp(\beta)$ and deduced the corresponding informative prior $\beta \sim N(0.8, 0.2^2)$. To complete the hierarchy, we have used $\alpha \sim Gamma(2, 0.1)$; $G_0 \sim N(\mu, \sigma^2)$, $\mu \sim N(0, 10)$ and $\sigma^{-2} \sim Gamma(4, 1)$ in all our simulations.

B.3: Construction of power priors for the DAMAT study data example

Asthma risk has been associated with $PM_{2.5}$/PM_{10} in many studies using both time-series and case-crossover designs. Among recent papers, an Alaskan study (Chimonas et al., 2007) found that a 10 $\mu g\, m^{-3}$ increase in PM_{10} was associated with a 0.6% (95% CI: 0.1%, 1.3%) increase in outpatient asthma visits, and a 1.8% (95% CI: 0.6%, 3.0%) increase in inhaled quick-relief mediation prescriptions. In Rio de Janeiro (Moura et al., 2009), a 10 $\mu g\, m^{-3}$ increase of PM_{10} was found to be associated with 6.7% (95% CI: 1.8%, 11.5%) increase for bronchial obstruction. In two Idaho cities (Ulirsch et al., 2007), a 24.3 $\mu g\, m^{-3}$ increase in PM_{10} was associated with 4.3% increase for respiratory disease. In the Detroit Medicaid population (Li et al., 2011), we found a 3-7% increase in asthma risks for a 9.2 $\mu g\, m^{-3}$ increase in $PM_{2.5}$. Larger effects were found when only the warmer season was considered (Villeneuve et al., 2007). These results are converted in terms of risk ratios in the following table. More detailed reviews can be found in Li et al. (2011).

Study	Risk Ratios*
Chimonas et al., 2007	1.006, 1.018
Moura et al., 2009	1.065
Ulirsch et al., 2007	1.017
Li et al., 2011	1.03-1.09

* Risk ratios $\exp(\hat{\beta}_{PM_{2.5}})$ and $\exp(\hat{\beta}_{PM_{10}})$ corresponding to 10 $\mu g\, m^{-3}$ increase in PM_x concentrations

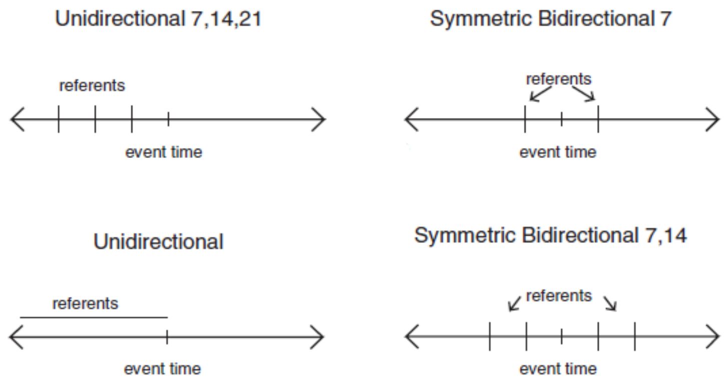
Based on these studies where different cohorts, statistical models, and variant asthma outcomes were used, we have a belief that the asthma-$PM_{2.5}$ association is in general modest with an odds ratio ranging (1.01-1.09) for a 10 $\mu g\, m^{-3}$ increase in $PM_{2.5}$ (if we assume that effect of PM_{10} has no substantial difference from that of $PM_{2.5}$). In our DAMAT data analysis section, we constructed a presumed 95% confidence interval (1.01,1.09) based on the above information, and took the prior mean to be the center ($\mu_\beta = [\log(1.09) + \log(1.01)]/2 = 0.05$) and the prior standard deviation to be one-fourth the width of the interval ($\sigma_\beta = [\log(1.09) - \log(1.01)]/4 = 0.02$), i.e., $\beta_{PM_{2.5}} \sim N(0.05, 0.02^2)$.

For the power priors, suppose we observed D_0 in terms of summary statistics from previous studies, e.g. the MLEs $\hat{\beta}_k$'s with variance $\hat{\sigma}^2_{\hat{\beta}_k}$'s. We assume the sampling distribution of $\hat{\beta}_k$ is normal, namely, $\hat{\beta}_k|\beta \sim N(\beta, \hat{\sigma}^2_{\hat{\beta}_k})$, $k = 1, ..., K$. Assuming the studies were independent and equally weighted, then

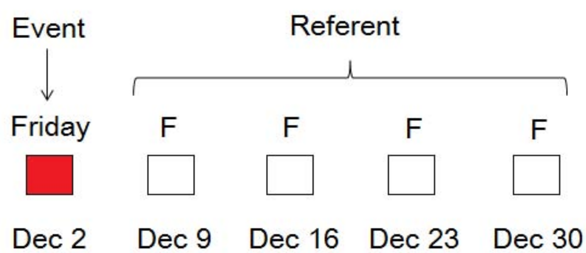
$$L(\beta|D_0) \propto \prod_{k=1}^{K} \exp(-\frac{(\beta - \hat{\beta}_k)^2}{2\hat{\sigma}^2_{\hat{\beta}_k}}).$$

In particular, we considered $K = 3$ prior studies having small, modest and strong effect sizes $(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3) = (0.02, 0.05, 0.08)$ with $(\hat{\sigma}^2_{\hat{\beta}_1}, \hat{\sigma}^2_{\hat{\beta}_2}, \hat{\sigma}^2_{\hat{\beta}_3}) = (0.02, 0.02, 0.03)$ respectively, to reflect $PM_{2.5}$-asthma association (change in asthma risk for a 10 $\mu g\, m^{-3}$ increase in $PM_{2.5}$) based on the evidence in a recent review paper (Li et al. 2011). $L(\beta|D_0)$ is then described by the product of the three independent normal likelihoods.

Web Appendix C: Supplementary figures and tables



(a) non-localizable designs



(b) time-stratified design

Appendix Figure 1. Referent time selections for case-crossover designs.

Appendix Table 1. Risk ratios of acute asthma events corresponding to a 10 $\mu g\, m^{-3}$ increase in $PM_{2.5}$ in the DAMAT study, using data in year 2006 only ($T = 365$).

	TSD[a]	
Frequentist	MLE[a]	95% CI[a]
Conditional likelihood	1.043	(0.983, 1.107)
TSLL[a]	1.043	(0.983, 1.107)
Full likelihood REM[a] (T)	1.041	(0.981, 1.105)
Bayesian (prior 1)[b]	Bayes[a]	95% HPD[a]
Conditional likelihood	1.042	(0.982, 1.095)
TSLL	1.042	(0.982, 1.096)
Full likelihood DP[b] (T)	1.040	(0.980, 1.093)
Bayesian (prior 2)	Bayes	95% HPD
Conditional likelihood	1.049	(1.017, 1.085)
TSLL	1.049	(1.018, 1.086)
Full likelihood DP (T)	1.050	(1.020, 1.088)
Bayesian (power prior 1)	Bayes	95% HPD
Conditional likelihood	1.040	(1.010, 1.068)
TSLL	1.040	(1.010, 1.068)
Full likelihood DP (T)	1.041	(1.011, 1.070)
Bayesian (power prior 2)	Bayes	95% HPD
Conditional likelihood	1.059	(1.019, 1.098)
TSLL	1.059	(1.020, 1.099)
Full likelihood DP (T)	1.060	(1.021, 1.102)
Bayesian (power prior 3)	Bayes	95% HPD
Conditional likelihood	1.040	(1.008, 1.075)
TSLL	1.040	(1.008, 1.074)
Full likelihood DP (T)	1.041	(1.010, 1.076)
Bayesian (power prior 4)	Bayes	95% HPD
Conditional likelihood	1.058	(1.007, 1.100)
TSLL	1.058	(1.008, 1.101)
Full likelihood DP (T)	1.060	(1.012, 1.106)

[a] TSD: time-stratified design; TSLL: time-stratified log-linear; REM: random effects model; MLE: maximum likelihood estimate (penalized pseudo-likelihood for REM); CI: confidence interval; Bayes: Bayes estimates in terms of posterior mean; HPD: highest posterior density.

[b] DP (T): Dirichlet process prior $DP(\alpha, G_0)$ on the random intercepts $\boldsymbol{\nu}$ in $L^T_{full}(\boldsymbol{\beta}, \boldsymbol{\nu})$ is used, where prior distribution $\alpha \sim Gamma(0.5, 0.1)$ is chosen. Prior 1: $\beta_{PM_{2.5}} \sim N(0, 10^2)$; Prior 2: $\beta_{PM_{2.5}} \sim N(0.05, 0.02^2)$; Power prior 1: $a_0 = 0.5$; Power prior 2: $a_0 = 1.0$; Power prior 3: $a_0 \sim Beta(20, 20)$ with mean 0.50 and variance 0.08; Power prior 4: $a_0 \sim Beta(50, 1)$ with mean 0.98 and variance 0.02.

Appendix Table 2. Posterior distributions derived under full likelihood $L^T_{full}(\boldsymbol{\beta}, \boldsymbol{\nu})$ with six different prior distributions on $\boldsymbol{\nu}$ as a sensitivity analysis, under a time-stratified case-crossover design for the DAMAT study.

	$\beta_{PM_{2.5}}$			α			Number of clusters		
Priors[b]	Mean[a]	SD[a]	Median[a]	Mean	SD	Median	Mean	SD	Median
prior A	0.059	0.016	0.058				1.00	0.00	1.00
prior B	0.058	0.016	0.058				238.64	13.22	238
prior G1	0.058	0.016	0.058	0.08	0.12	0.04	1.09	0.33	1
prior G2	0.058	0.015	0.057	0.34	0.27	0.28	1.43	0.80	1
prior G3	0.058	0.016	0.058	0.08	0.11	0.04	1.06	0.27	1
prior G4	0.058	0.016	0.058	5.62	1.82	5.34	15.50	7.45	14

[a] posterior mean, standard deviation and median.

[b] Dirichlet process prior $DP(\alpha, G_0)$ on $\boldsymbol{\nu}$ in $L^T_{full}(\boldsymbol{\beta}, \boldsymbol{\nu})$. The base prior setting on G_0 was used as described in section 6; the priors on α was varied as follows.
prior A: $\nu_t = \nu^*$ for $t = 1, ..., 1096$, where $\nu^* \sim N(0, 10^2)$; prior B: $\nu_t \overset{iid}{\sim} N(0, 10^2)$;
prior G1: $\alpha \sim Gamma(0.5, 0.1)$; prior G2: $\alpha \sim Gamma(2, 0.2)$;
prior G3: $\alpha \sim Gamma(10, 0.5)$; prior G4: $\alpha \sim Gamma(20, 1)$.

Additional References:

1. Moura M, Junger WL, Mendonca GA, et al. (2009). Air quality and emergency pediatric care for symptoms of bronchial obstruction categorized by age bracket in Rio De Janeiro, Brazil. Cad Saude Publica. 25(3), 635-44.

2. Chimonas MA, Gessner BD. (2007). Airborne particulate matter from primarily geologic, non-industrial sources at levels below national ambient air quality standards is associated with outpatient visits for asthma and quick-relief medication prescriptions among children less than 20 years old enrolled in Medicaid in Anchorage, Alaska. Environ Res. 103, 397-404.

3. Ulirsch GV, Ball LM, Kaye W, et al. (2007). Effect of particulate matter air pollution on hospital admissions and medical visits for lung and heart disease in two southeast Idaho Cities. J Expo Sci Environ Epidemiol. 17, 478-487.

4. Villeneuve P, Chen L, Rowe BH, et al. (2007). Outdoor air pollution and emergency department visits among children and adults: a case-crossover study in northern Alberta, Canada. Environ Health. 6(40).