## APPENDIX

### 1. Details of the power calculation for the case-crossover-model

For each event day i, let $x_{i1}$ denote the value of the pollutant variable on that day, and let $x_{i0}$ and $x_{i2}$ denote the values of the pollutant variable on the two control days (e.g., one week before and one week after the event day).

The conditional probability of the event happening on the event day i given that an event occurs on one of the three days is then usually modeled as

$$\frac{\exp(\beta\cdot x_{i1}+\eta_{i1})}{1+\exp(\beta\cdot x_{i1}+\eta_{i1})}\frac{1}{1+\exp(\beta\cdot x_{i0}+\eta_{i0})}\frac{1}{1+\exp(\beta\cdot x_{i2}+\eta_{i2})}\Big/\sum_{j=0}^{2}\frac{\exp(\beta\cdot x_{ij}+\eta_{ij})}{1+\exp(\beta\cdot x_{i1}+\eta_{i1})}\frac{1}{1+\exp(\beta\cdot x_{i0}+\eta_{i0})}\frac{1}{1+\exp(\beta\cdot x_{i2}+\eta_{i2})}$$

.

This expression can be simplified to

$$\exp(\beta\cdot x_{i1}+\eta_{i1})\Big/\sum_{j=0}^{2}\exp(\beta\cdot x_{ij}+\eta_{ij})\,.$$

If we assume that the terms $\eta_{ij}$ (comprising the effects of other factors and the intercept parameter) are (essentially) constant over the time interval between the two control days of an event (assumption 1), then this may be further simplified to

$$\exp(\beta\cdot x_{i1})\Big/\sum_{j=0}^{2}\exp(\beta\cdot x_{ij})\,.$$

The likelihood function for the total sample is then given by

$$L(\beta)=\prod_{i=1}^{n}\frac{\exp(\beta\cdot x_{i1})}{\exp(\beta\cdot x_{i0})+\exp(\beta\cdot x_{i1})+\exp(\beta\cdot x_{i2})}=\prod_{i=1}^{n}\frac{1}{\exp(\beta\cdot(x_{i0}-x_{i1}))+1+\exp(\beta\cdot(x_{i2}-x_{i1}))}\,,$$

and the corresponding log-likelihood function by

$$l(\beta)\;=\;-\sum_{i=1}^{n}\log[\exp(\beta\cdot(x_{i0}-x_{i1}))+1+\exp(\beta\cdot(x_{i2}-x_{i1}))]\,.$$

The equation for the estimate $\hat{\beta}$ of $\beta$ is then given by

$$l'(\hat{\beta})\;=\;-\sum_{i=1}^{n}\frac{(x_{i0}-x_{i1})\exp(\hat{\beta}\cdot(x_{i0}-x_{i1}))+(x_{i2}-x_{i1})\exp(\hat{\beta}(x_{i2}-x_{i1}))}{\exp(\hat{\beta}\cdot(x_{i0}-x_{i1}))+1+\exp(\hat{\beta}\cdot(x_{i2}-x_{i1}))}=0$$

And the asymptotic variance of $\hat{\beta}$ by $-l''(\beta)^{-1}$.

Now, we have

$$l''(\beta)\;=\;-\sum_{i=1}^{n}\frac{[(x_{i0}-x_{i1})^2\exp(\beta\cdot(x_{i0}-x_{i1}))+(x_{i2}-x_{i1})^2\exp(\beta\cdot(x_{i2}-x_{i1}))]D_i-[(x_{i0}-x_{i1})\exp(\beta\cdot(x_{i0}-x_{i1}))+(x_{i2}-x_{i1})\exp(\beta\cdot(x_{i2}-x_{i1}))]^2}{D_i^2}$$

where $D_i = \exp(\beta \cdot (x_{i0} - x_{i1})) + 1 + \exp(\beta \cdot (x_{i2} - x_{i1}))$ .

If the terms $\beta \cdot (x_{ij} - x_{i1})$ are all close to 0, as it is the case if the effects from x are small (assumption 2), then $D_i \approx 3$, for all i and

$$l''(\beta) \approx -\sum_{i=1}^{n} \frac{3[(x_{i0}-x_{i1})^2 + (x_{i2}-x_{i1})^2] - [(x_{i0}-x_{i1}) + (x_{i2}-x_{i1})]^2}{9}$$
$$= -\sum_{i=1}^{n} \frac{2[(x_{i0}-x_{i1})^2 + (x_{i2}-x_{i1})^2 - (x_{i0}-x_{i1})(x_{i2}-x_{i1})]}{9}$$
$$= -\sum_{i=1}^{n} \frac{(x_{i0}-x_{i1})^2 + (x_{i2}-x_{i1})^2 + (x_{i2}-x_{i0})^2}{9}$$

Under the two assumptions stated, the ratio between the asymptotic variance of $\hat{\beta}$ and

$$9 / \sum_{i=1}^{n} (x_{i0}-x_{i1})^2 + (x_{i2}-x_{i1})^2 + (x_{i2}-x_{i0})^2$$

is thus approximately equal to 1.
Then we have

$$\text{variance of } \hat{\beta} \approx \left( (\text{number of events}) \times \frac{V(x_{i0}-x_{i1}) + V(x_{i2}-x_{i1}) + V(x_{i2}-x_{i0})}{9} \right)^{-1},$$

where $V(d_i)$ denotes the sample variance of the respective difference $d_i$ , and

$$\text{standard error of } \hat{\beta} \approx$$
$$3\left((\text{number of events}) \times [V(x_{i0}-x_{i1}) + V(x_{i2}-x_{i1}) + V(x_{i2}-x_{i0})]\right)^{-1/2}.$$

**2. Details of the power calculation for the Poisson-regression model**

Let $y_i$ denote the number of events on day i and let $x_i$ denote the value of the pollutant variable of interest on that same day. Moreover, we assume that the count data $y_i$ are regressed against splines of time and other time-dependent covariates (e.g., meteorological variables) using a Poisson model and that the pollutant data $x_i$ are regressed against the same splines and covariates using ordinary least squares. Then we obtain predicted values $\hat{y}_i$ and $\hat{x}_i$ and can consider the residuals $dx_i = x_i - \hat{x}_i$ . The "residual" (or short-term) effect of $x_i$ on $y_i$ after adjusting for seasonality (modeled by the temporal splines) and the other time-dependent covariates may then be estimated through the equation

$$0 = \sum_{i=1}^{n} dx_i \, [y_i - \exp(\log(\hat{y}_i) + \beta \, dx_i)] .$$

If the effects from $x_i$ on $y_i$ are small, then this estimate is close to the one obtained when all the splines and covariates are simultaneously included in the Poisson regression model along with $x_i$.

The asymptotic variance of the above estimate $\hat{\beta}$ is then given by the inverse of

$$\sum_{i=1}^{n} dx_i^2 \, \exp(\log(\hat{y}_i) + \beta \, dx_i) ,$$

provided that there is no overdispersion.

If we again assume that the terms $\beta dx_i$ are close to 0, then the asymptotic variance of $\hat{\beta}$ can be approximated by $(\sum_{i=1}^{n} dx_i^2\ \hat{y}_i)^{-1}$.

If $E(dx_i^2)$ varies little with time and short term variation in $\hat{y}_i$ is small, then this expression can be approximated by $(\frac{1}{n}\sum_{i=1}^{n} dx_i^2 \sum_{i=1}^{n} \hat{y}_i)^{-1}$.

On the other hand, the predicted values $\hat{y}_i$ satisfy $\sum_{i=1}^{n} \hat{y}_i = \sum_{i=1}^{n} y_i = \text{total number of events}$.

Thus, under these conditions, we have

$$\text{variance of } \hat{\beta} \approx \left((\text{total number of events}) \times (\text{variance of } \mathrm{dx_i})\right)^{-1},$$

and

$$\text{standard error of } \hat{\beta} \approx \left((\text{total number of events}) \times (\text{variance of } \mathrm{dx_i})\right)^{-1/2}$$