Web-based Supplementary Materials for
“A space-time conditional intensity model for invasive meningococcal disease occurence”

by

Sebastian Meyer[1,2], Johannes Elias[3], and Michael Höhle[4,2]

[1]Department of Psychiatry and Psychotherapy, Ludwig-Maximilians-Universität, München, Germany

[2]Department of Statistics, Ludwig-Maximilians-Universität, München, Germany

[3]German Reference Centre for Meningococci, University of Würzburg, Würzburg, Germany

[4]Department for Infectious Disease Epidemiology, Robert Koch Institute, Berlin, Germany

August 3, 2011

Web Appendix A: Calculus of the Score Function

Let $\boldsymbol{\vartheta}$ denote any subvector of $\boldsymbol{\theta}$. Then, the partial derivative of the log-likelihood with respect to $\boldsymbol{\vartheta}$ is

$$\boldsymbol{s}_{\boldsymbol{\vartheta}}(\boldsymbol{\theta}) := \frac{\partial}{\partial \boldsymbol{\vartheta}} l(\boldsymbol{\theta}) = \sum_{i=1}^{n} \frac{\frac{\partial}{\partial \boldsymbol{\vartheta}} \lambda^*_{\boldsymbol{\theta}}(t_i, \boldsymbol{s}_i, \kappa_i)}{\lambda^*_{\boldsymbol{\theta}}(t_i, \boldsymbol{s}_i, \kappa_i)} - \int_0^T \int_W \sum_{\kappa \in \mathcal{K}} \frac{\partial}{\partial \boldsymbol{\vartheta}} \lambda^*_{\boldsymbol{\theta}}(t, \boldsymbol{s}, \kappa) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{s} \, , \tag{1}$$

and the score function is $\boldsymbol{s}(\boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} l(\boldsymbol{\theta}) = \left(\boldsymbol{s}'_{\boldsymbol{\beta_0}}, \boldsymbol{s}'_{\boldsymbol{\beta}}, \boldsymbol{s}'_{\boldsymbol{\gamma}}, \boldsymbol{s}'_{\boldsymbol{\sigma}}, \boldsymbol{s}'_{\boldsymbol{\alpha}}\right)' (\boldsymbol{\theta})$. The necessary partial derivatives of the CIF with their respective time-space-mark integrals are given in the following subsections and can be plugged into the equation (1). The analytic derivatives of the interaction function f and g with respect to $\boldsymbol{\sigma}$ and $\boldsymbol{\alpha}$, respectively, have to be determined for the specific model at hand. For instance, a type-specific spatial Gaussian kernel

$$f_{\boldsymbol{\sigma}}(\boldsymbol{s}|\kappa) = \exp\left(-\frac{\|\boldsymbol{s}\|^2}{2\sigma_\kappa^2}\right)$$

with $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_K)'$ has partial derivatives

$$\frac{\partial}{\partial \sigma_k} f_{\boldsymbol{\sigma}}(\boldsymbol{s}|\kappa) = \mathbb{1}_{k=\kappa}(\kappa) \cdot \exp\left(-\frac{\|\boldsymbol{s}\|^2}{2\sigma_k^2}\right) \frac{\|\boldsymbol{s}\|^2}{\sigma_k^3}, \qquad \text{for any } k \in \mathcal{K} \, .$$

The type-specific temporal interaction function $g_{\boldsymbol{\alpha}}(t|\kappa) = e^{-\alpha_\kappa t}$ with $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)$ has partial derivatives $\frac{\partial}{\partial \alpha_k} g_{\boldsymbol{\alpha}}(t|\kappa) = \mathbb{1}_{k=\kappa}(\kappa) \cdot (-t \, e^{-\alpha_k t})$, for any $k \in \mathcal{K}$. While the integral of $\frac{\partial}{\partial \sigma_\kappa} f_{\boldsymbol{\sigma}}(\boldsymbol{s}|\kappa)$ over the region R_j will be approximated by numerical integration, the temporal function $\frac{\partial}{\partial \alpha_\kappa} g_{\boldsymbol{\alpha}}(t|\kappa)$ is assumed to permit analytical integration.

Endemic intercept(s) $\boldsymbol{\beta_0}$:

Let $\beta_{0,k}$, $k \in \{1,\dots,K\}$ be one of the type-specific intercepts in $\boldsymbol{\beta_0}$. Then,

$$\frac{\partial}{\partial \beta_{0,k}} \lambda^*_{\boldsymbol{\theta}}(t,\boldsymbol{s},\kappa) = \mathbb{1}_{k=\kappa}(\kappa) \cdot \exp\left(\beta_{0,k} + o_{\xi(\boldsymbol{s})} + \boldsymbol{\beta}' \boldsymbol{z}_{\tau(t),\xi(\boldsymbol{s})}\right)$$

since the parameter $\beta_{0,k}$ appears in the endemic component $h_{\boldsymbol{\theta}}(t,\boldsymbol{s},\kappa)$ if and only if $\kappa = k$. The corresponding integrated value is

$$\int_0^T \int_W \sum_{\kappa\in\mathcal{K}} \frac{\partial}{\partial \beta_{0,k}} \lambda^*_{\boldsymbol{\theta}}(t,\boldsymbol{s},\kappa) \,\mathrm{d}t\,\mathrm{d}\boldsymbol{s} = e^{\beta_{0,k}} \cdot \sum_{\tau=1}^{D}\sum_{\xi=1}^{M} |C_\tau||A_\xi| \exp\left(o_\xi + \boldsymbol{\beta}'\boldsymbol{z}_{\tau,\xi}\right) ,$$

cf. the integral of the endemic component in equation (6) of the paper. If the model assumes a type-invariant endemic intercept $\boldsymbol{\beta_0} = \beta_0$, then

$$\frac{\partial}{\partial \beta_0} \lambda^*_{\boldsymbol{\theta}}(t,\boldsymbol{s},\kappa) = \exp\left(\beta_0 + o_{\xi(\boldsymbol{s})} + \boldsymbol{\beta}' \boldsymbol{z}_{\tau(t),\xi(\boldsymbol{s})}\right)$$

with integrated value

$$\int_0^T \int_W \sum_{\kappa\in\mathcal{K}} \frac{\partial}{\partial \beta_0} \lambda^*_{\boldsymbol{\theta}}(t,\boldsymbol{s},\kappa) \,\mathrm{d}t\,\mathrm{d}\boldsymbol{s} = K\, e^{\beta_0} \cdot \sum_{\tau=1}^{D}\sum_{\xi=1}^{M} |C_\tau||A_\xi| \exp\left(o_\xi + \boldsymbol{\beta}'\boldsymbol{z}_{\tau,\xi}\right) .$$

Endemic covariate effects $\boldsymbol{\beta}$:

$$\frac{\partial}{\partial \boldsymbol{\beta}} \lambda^*_{\boldsymbol{\theta}}(t,\boldsymbol{s},\kappa) = \exp\left(\beta_0^(\kappa) + o_{\xi(\boldsymbol{s})} + \boldsymbol{\beta}' \boldsymbol{z}_{\tau(t),\xi(\boldsymbol{s})}\right) \cdot \boldsymbol{z}_{\tau(t),\xi(\boldsymbol{s})}$$

with corresponding integral vector (element-wise integral values)

$$\left(\sum_{\kappa\in\mathcal{K}} \exp\left(\beta_0(\kappa)\right)\right) \cdot \sum_{\tau=1}^{D}\sum_{\xi=1}^{M} |C_\tau||A_\xi| \exp(o_\xi + \boldsymbol{\beta}'\boldsymbol{z}_{\tau,\xi})\, \boldsymbol{z}_{\tau,\xi} .$$

Epidemic effects γ:

$$\frac{\partial}{\partial \boldsymbol{\gamma}} \lambda^*_{\boldsymbol{\theta}}(t, \boldsymbol{s}, \kappa) = \sum_{j \in I^*(t,\boldsymbol{s},\kappa)} e^{\boldsymbol{\gamma}' \boldsymbol{m}_j} \, g_{\boldsymbol{\alpha}}(t - t_j|\kappa_j) \, f_{\boldsymbol{\sigma}}(\boldsymbol{s} - \boldsymbol{s}_j|\kappa_j) \, \boldsymbol{m}_j \, ,$$

and the corresponding integral can be deduced similar to equation (7) of the paper as

$$\sum_{j=1}^{n} q_{\kappa_j,\bullet} \, e^{\boldsymbol{\gamma}' \boldsymbol{m}_j} \left[\int_0^{\min\{T - t_j;\varepsilon\}} g_{\boldsymbol{\alpha}}(t|\kappa_j) \, \mathrm{d}t \right] \left[\int_{R_j} f_{\boldsymbol{\sigma}}(\boldsymbol{s}|\kappa_j) \, \mathrm{d}\boldsymbol{s} \right] \boldsymbol{m}_j \, .$$

Parameters σ and α of the interaction functions:

For a general spatial kernel $f_{\boldsymbol{\sigma}}(\boldsymbol{s}|\kappa)$,

$$\frac{\partial}{\partial \boldsymbol{\sigma}} \lambda^*_{\boldsymbol{\theta}}(t, \boldsymbol{s}, \kappa) = \sum_{j \in I^*(t,\boldsymbol{s},\kappa)} e^{\eta_j} \, g_{\boldsymbol{\alpha}}(t - t_j|\kappa_j) \, \left[\frac{\partial}{\partial \boldsymbol{\sigma}} f_{\boldsymbol{\sigma}}(\boldsymbol{s} - \boldsymbol{s}_j|\kappa_j) \right]$$

with corresponding integral

$$\sum_{j=1}^{n} q_{\kappa_j,\bullet} \, e^{\eta_j} \left[\int_0^{\min\{T - t_j;\varepsilon\}} g_{\boldsymbol{\alpha}}(t|\kappa_j) \, \mathrm{d}t \right] \left[\int_{R_j} \frac{\partial}{\partial \boldsymbol{\sigma}} f_{\boldsymbol{\sigma}}(\boldsymbol{s}|\kappa_j) \, \mathrm{d}\boldsymbol{s} \right] .$$

Similarly, for a general temporal kernel $g_{\boldsymbol{\alpha}}(t|\kappa)$,

$$\frac{\partial}{\partial \boldsymbol{\alpha}} \lambda^*_{\boldsymbol{\theta}}(t, \boldsymbol{s}, \kappa) = \sum_{j \in I^*(t,\boldsymbol{s},\kappa)} e^{\eta_j} \, \left[\frac{\partial}{\partial \boldsymbol{\alpha}} g_{\boldsymbol{\alpha}}(t - t_j|\kappa_j) \right] \, f_{\boldsymbol{\sigma}}(\boldsymbol{s} - \boldsymbol{s}_j|\kappa_j)$$

with corresponding integral

$$\sum_{j=1}^{n} q_{\kappa_j,\bullet} \, e^{\eta_j} \left[\int_0^{\min\{T - t_j;\varepsilon\}} \frac{\partial}{\partial \boldsymbol{\alpha}} g_{\boldsymbol{\alpha}}(t|\kappa_j) \, \mathrm{d}t \right] \left[\int_{R_j} f_{\boldsymbol{\sigma}}(\boldsymbol{s}|\kappa_j) \, \mathrm{d}\boldsymbol{s} \right] .$$

Web Appendix B: Fisher Information Matrix

The inverse of the Fisher information matrix at the maximum likelihood estimate (MLE) $\hat{\boldsymbol{\theta}}_{ML}$ is in general likelihood theory used as an estimate of the variance of $\hat{\boldsymbol{\theta}}_{ML}$. The precise conditions under which asymptotic properties of MLEs hold for spatio-temporal point processes have been established by Rathbun (1996). Specifically, the conditions for existence, consistence and asymptotic normality of a local maximum $\hat{\boldsymbol{\theta}}_{ML}$ as $T \to \infty$ for a fixed observation region W are discussed in Meyer (2009, Section 4.2.3).

The expected Fisher information $\mathcal{I}(\boldsymbol{\theta})$ can be estimated by the "optional variation process" adapted to the marked spatio-temporal setting (Rathbun, 1996, equation (4.7))

$$\int_0^T \int_W \int_{\mathcal{K}} \left(\frac{\partial}{\partial \boldsymbol{\theta}} \log \lambda^*_{\boldsymbol{\theta}}(t, \boldsymbol{s}, \kappa) \right)^{\otimes 2} \, \mathrm{d}N(t, \boldsymbol{s}, \kappa)$$

through its observed realisation

$$\hat{\mathcal{I}}(\boldsymbol{\theta}) = \sum_{i=1}^{n} \left(\frac{\partial}{\partial \tilde{\boldsymbol{\theta}}} \log \lambda^*_{\tilde{\boldsymbol{\theta}}}(t_i, \boldsymbol{s}_i, \kappa_i) \bigg|_{\tilde{\boldsymbol{\theta}}=\boldsymbol{\theta}} \right)^{\otimes 2} = \sum_{i=1}^{n} \left(\frac{\frac{\partial}{\partial \tilde{\boldsymbol{\theta}}} \lambda^*_{\tilde{\boldsymbol{\theta}}}(t_i, \boldsymbol{s}_i, \kappa_i)}{\lambda^*_{\tilde{\boldsymbol{\theta}}}(t_i, \boldsymbol{s}_i, \kappa_i)} \Bigg|_{\tilde{\boldsymbol{\theta}}=\boldsymbol{\theta}} \right)^{\otimes 2} ,$$

where $\boldsymbol{a}^{\otimes 2} := \boldsymbol{a}\boldsymbol{a}'$ for a vector $\boldsymbol{a}$. Uncertainty of the parameter estimates is thus deduced from the diagonal of $\hat{\mathcal{I}}^{-1/2}(\hat{\boldsymbol{\theta}}_{ML})$, which contains their standard errors.

References

Meyer, S. (2009). Spatio-temporal infectious disease epidemiology based on point processes. Master's thesis, Department of Statistics, Ludwig-Maximilians-Universität, München. Available as `http://epub.ub.uni-muenchen.de/11703/`.

Rathbun, S. L. (1996). Asymptotic properties of the maximum likelihood estimator for spatio-temporal point processes. *Journal of Statistical Planning and Inference* **51,** 55–74.

Web Appendix C: Simulation Algorithm

This appendix provides a more implementational view on the simulation algorithm described in Section 4 of the paper. In addition to the notation of the paper, let $L(t)$ be the next time point after time t where any endemic covariate in any tile changes its value, or a previously infected individual stops spreading the disease, i.e.:

$$L(t) = \min\left\{t' > t \mid (\exists\, \xi \in \{1,\ldots,M\} : \boldsymbol{z}_{\tau(t'),\xi} \neq \boldsymbol{z}_{\tau(t),\xi}) \vee (\exists\, j \in \{1,\ldots,N_g(t)\} : t' = t_j + \varepsilon)\right\}.$$

An implementational perspective on the simulation algorithm is then:

Algorithm 1: Ogata's modified thinning adapted for `twinstim`

1 Given current time t, update $L(t)$ and calculate local upper bound $\overline{\lambda_g^*}(t)$;
2 Generate proposed waiting time $\Delta \sim \operatorname{Exp}(\overline{\lambda_g^*}(t))$;
3 **if** $t + \Delta > L(t)$ **then**
4   Let $t = L(t)$;
5 **else**
6   Let $t = t + \Delta$;
7   Accept t with probability $\lambda_g^*(t)/\overline{\lambda_g^*}(t)$, otherwise goto step 1;
8   Draw the source of infection (infective individual j or endemic) with weights equal to the respective components of $\lambda_g^*(t)$ (cf. equation (8) of the paper);
9   **if** *endemic source of infection* **then**
10   Draw the type κ of the new event with weights $\exp(\beta_0(\kappa))$, $\kappa \in \mathcal{K}$;
11   Draw the tile A_ξ of the new event with weights $|A_\xi|\, \rho_{\tau(t),\xi}\, e^{\boldsymbol{\beta}' \boldsymbol{z}_{\tau(t),\xi}}$, $\xi \in \{1,\ldots,M\}$;
12   Draw the location $\boldsymbol{s}$ of the new event uniform within the sampled tile A_ξ;
13   **else**
14   Draw the type κ of the new event at random out of the types which can be triggered by the source individual j, i.e. draw from $U(\{\kappa \in \mathcal{K} : q_{\kappa_j,\kappa} = 1\})$;
15   Draw the relative location $\boldsymbol{v}$ of the new event (relative to the source j) from the density $f(\boldsymbol{s}|\kappa_j) / \int_{R_j} f(\boldsymbol{s}|\kappa_j)\, \mathrm{d}s$ on R_j, i.e. $\boldsymbol{s} = \boldsymbol{s}_j + \boldsymbol{v}$;
16   Draw additional marks according to the pre-specified distribution;
17   Update the event history;
18 Goto step 1;

An `R` implementation of the algorithm can be found as the function `simEpidataCS` in the popular `surveillance` package.