Geostatistical modelling of the relationship between microfilariae and antigenaemia prevalence of lymphatic filariasis infection

Emanuele Giorgi[1], Jorge Cano[2], Rachel Pullan[2]

[1] Lancaster Medical School, Lancaster University, Lancaster, UK
[2] London School of Hygiene and Tropical Medicine, London, UK

RSS 2016 International Conference, 5-8 September, University of Manchester

Overview

- Lymphatic filariasis: what is it? What diagnostics?
- Bivariate geostatistical modelling of prevalence from two different diagnostics.
  1. A semi-mechanistic model for lymphatic filariasis microfilariae and antigenaemia prevalence.
  2. An empirical model for prevalence from any two diagnostics.
- Application to lymphatic filariasis prevalence data from West Africa.
- Discussion.

Lymphatic filariasis: the disease

Figure 1: Microfilaria of *Wuchereria*.

Figure 2: Microfilaria of *Brugia malayi*.

Figure 3: Patient with lymphedema.

Lymphatic filariasis: the disease

Figure 4: Endemic areas for LF in red.

Lymphatic filariasis: the vector

Figure 5: *Anopheles*.

Figure 6: *Culex*.

Figure 7: *Aedes*.

Lymphatic filariasis: the life cycle

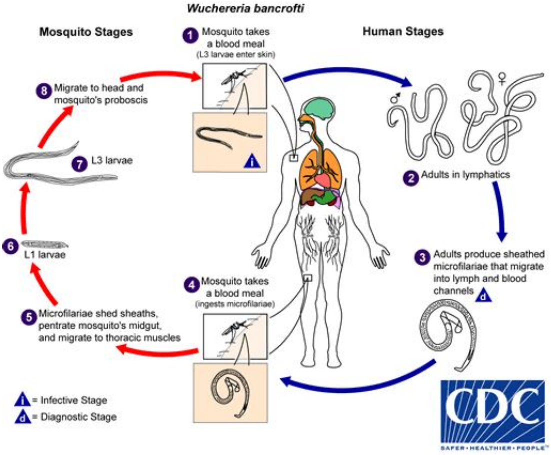


Figure 8: Life Cycle of *Wuchereria bancrofti*.

Lymphatic filariasis: diagnosis

Figure 9: Counting microfilariae at night.

Figure 10: ICT card for LF antigens detection.

Research question

- The data

$$\begin{aligned}\mathcal{D}_{MF} &= \{(x_{i,1}, n_{i,1}, y_{i,1}) : x_{i,1} \in A\},\\ \mathcal{D}_{ICT} &= \{(x_{i,2}, n_{i,2}, y_{i,2}) : x_{i,2} \in A\}.\end{aligned}$$

Research question

- The data

$$\begin{aligned}\mathcal{D}_{MF} &= \{(x_{i,1}, n_{i,1}, y_{i,1}) : x_{i,1} \in A\},\\ \mathcal{D}_{ICT} &= \{(x_{i,2}, n_{i,2}, y_{i,2}) : x_{i,2} \in A\}.\end{aligned}$$

- A model for the data

$$Y_{i,j} | S_j(x_{i,j}), U_{i,j} \sim \text{Binomial}(n_{i,j}, p_j(x_{i,j})), i = 1, \ldots, n_j, j = 1, 2.$$

Research question

- The data

$$\begin{aligned}\mathcal{D}_{MF} &= \{(x_{i,1}, n_{i,1}, y_{i,1}) : x_{i,1} \in A\},\\ \mathcal{D}_{ICT} &= \{(x_{i,2}, n_{i,2}, y_{i,2}) : x_{i,2} \in A\}.\end{aligned}$$

- A model for the data

$$Y_{i,j} | S_j(x_{i,j}), U_{i,j} \sim \text{Binomial}(n_{i,j}, p_j(x_{i,j})), i = 1, \ldots, n_j, j = 1, 2.$$

Objective

How should we build a bivariate geostatistical model for $p_1(x)$ and $p_2(x)$?

A semi-mechanistic approach (1)

A semi-mechanistic approach (1)

- W = ``number of worms in a sampled individual''.

A semi-mechanistic approach (1)

- W = ``number of worms in a sampled individual''.
- M = ``MF counts''.

A semi-mechanistic approach (1)

- W = ``number of worms in a sampled individual''.
- M = ``MF counts''.
- **Assumptions.** $W \sim$ Poisson(λ), $M|W = w \sim$ Poisson(rw).

A semi-mechanistic approach (1)

- W = ``number of worms in a sampled individual''.
- M = ``MF counts''.
- **Assumptions.** $W \sim$ Poisson(λ), $M|W = w \sim$ Poisson(rw).
- ICT: $p_1 = P(W > 0) = \phi(1 - exp\{-\lambda\})$. ($\phi = 0.97$)

A semi-mechanistic approach (1)

- W = ``number of worms in a sampled individual''.
- M = ``MF counts''.
- **Assumptions.** $W \sim$ Poisson(λ), $M|W = w \sim$ Poisson(rw).
- ICT: $p_1 = P(W > 0) = \phi(1 - exp\{-\lambda\})$. ($\phi = 0.97$)
- MF:

$$p_2 = P(M > 0) \quad = \quad 1 - P(M = 0)$$

A semi-mechanistic approach (1)

- W = ``number of worms in a sampled individual''.
- M = ``MF counts''.
- **Assumptions.** $W \sim$ Poisson(λ), $M|W = w \sim$ Poisson(rw).
- ICT: $p_1 = P(W > 0) = \phi(1 - exp\{-\lambda\})$. ($\phi = 0.97$)
- MF:

$$
\begin{aligned}
p_2 = P(M > 0) &= 1 - P(M = 0) \\
&= 1 - \sum_{w=0}^{+\infty} P(M = m|W = w)P(W = w)
\end{aligned}
$$

A semi-mechanistic approach (1)

- W = ``number of worms in a sampled individual''.
- M = ``MF counts''.
- **Assumptions.** $W \sim$ Poisson(λ), $M|W = w \sim$ Poisson(rw).
- ICT: $p_1 = P(W > 0) = \phi(1 - exp\{-\lambda\})$. ($\phi = 0.97$)
- MF:

$$
\begin{aligned}
p_2 = P(M > 0) &= 1 - P(M = 0) \\
&= 1 - \sum_{w=0}^{+\infty} P(M = m|W = w)P(W = w) \\
&= 1 - \exp[-r(1 - \exp\{-\lambda\})]
\end{aligned}
$$

A semi-mechanistic approach (1)

- W = ``number of worms in a sampled individual''.
- M = ``MF counts''.
- **Assumptions.** $W \sim$ Poisson(λ), $M|W = w \sim$ Poisson(rw).
- ICT: $p_1 = P(W > 0) = \phi(1 - exp\{-\lambda\})$. ($\phi = 0.97$)
- MF:

$$
\begin{aligned}
p_2 = P(M > 0) &= 1 - P(M = 0) \\
&= 1 - \sum_{w=0}^{+\infty} P(M = m|W = w)P(W = w) \\
&= 1 - \exp[-r(1 - \exp\{-\lambda\})] \\
&= 1 - \exp[-r\phi^{-1}p_1]
\end{aligned}
$$

A semi-mechanistic approach (2)

- What varies spatially?

A semi-mechanistic approach (2)

- What varies spatially?
- Density-independence: $\lambda(x)$ and $r(x) = \alpha > 0$ for all x.

A semi-mechanistic approach (2)

- What varies spatially?
- Density-independence: $\lambda(x)$ and $r(x) = \alpha > 0$ for all x.
- Density-dependence: $\lambda(x)$ and $r(x) = \alpha\lambda(x)^{\gamma}$, $\gamma \in \mathbb{R}$.

A semi-mechanistic approach (2)

- What varies spatially?
- Density-independence: $\lambda(x)$ and $r(x) = \alpha > 0$ for all x.
- Density-dependence: $\lambda(x)$ and $r(x) = \alpha\lambda(x)^{\gamma}$, $\gamma \in \mathbb{R}$.
- $\log\{\lambda(x)\} = d(x)^{\top}\beta + S(x) + Z(x)$.

A semi-mechanistic approach (2)

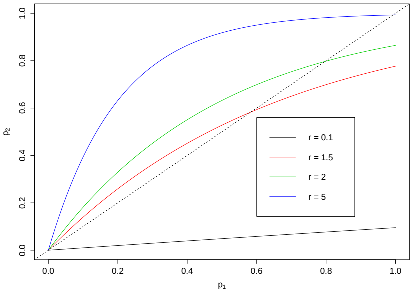
- What varies spatially?
- Density-independence: $\lambda(x)$ and $r(x) = \alpha > 0$ for all x.
- Density-dependence: $\lambda(x)$ and $r(x) = \alpha\lambda(x)^{\gamma}$, $\gamma \in \mathbb{R}$.
- $\log\{\lambda(x)\} = d(x)^{\top}\beta + S(x) + Z(x)$.
- $S(x) \sim \text{GP}(0, \sigma^2, \rho(\cdot;\phi))$.

A semi-mechanistic approach (2)

- What varies spatially?
- Density-independence: $\lambda(x)$ and $r(x) = \alpha > 0$ for all x.
- Density-dependence: $\lambda(x)$ and $r(x) = \alpha\lambda(x)^{\gamma}$, $\gamma \in \mathbb{R}$.
- $\log\{\lambda(x)\} = d(x)^{\top}\beta + S(x) + Z(x)$.
- $S(x) \sim \text{GP}(0, \sigma^2, \rho(\cdot;\phi))$.
- $Z(x) \sim N(0, \tau^2)$ i.i.d.

A semi-mechanistic approach (2)

- What varies spatially?
- Density-independence: $\lambda(x)$ and $r(x) = \alpha > 0$ for all x.
- Density-dependence: $\lambda(x)$ and $r(x) = \alpha\lambda(x)^{\gamma}$, $\gamma \in \mathbb{R}$.
- $\log\{\lambda(x)\} = d(x)^{\top}\beta + S(x) + Z(x)$.
- $S(x) \sim$ GP$(0, \sigma^2, \rho(\cdot;\phi))$.
- $Z(x) \sim N(0, \tau^2)$ i.i.d.



An empirical approach (1)

An empirical approach (1)

- ICT: $\log\{p_1(x)/[1-p_1(x)]\} = d_1(x)^\top \beta_1 + S_1(x) + Z_1(x)$.

An empirical approach (1)

- ICT: $\log\{p_1(x)/[1 - p_1(x)]\} = d_1(x)^\top \beta_1 + S_1(x) + Z_1(x)$.
- MF:

$$\begin{aligned}\log\{p_2(x)/[1 - p_2(x)]\} &= d_2(x)^\top \beta_2 + S_2(x) + Z_2(x) + \\ &\quad \gamma f(p_1(x))\end{aligned}$$

where $f : [0, 1] \to I \subseteq \mathbb{R}$ continuous non-decreasing.

An empirical approach (1)

- ICT: $\log\{p_1(x)/[1-p_1(x)]\} = d_1(x)^\top \beta_1 + S_1(x) + Z_1(x)$.
- MF:

$$\log\{p_2(x)/[1-p_2(x)]\} = d_2(x)^\top \beta_2 + S_2(x) + Z_2(x) + \gamma f(p_1(x))$$

  where $f : [0,1] \to I \subseteq \mathbb{R}$ continuous non-decreasing.
- Separate models if $\gamma = 0$.

An empirical approach (1)

- ICT: $\log\{p_1(x)/[1-p_1(x)]\} = d_1(x)^\top\beta_1 + S_1(x) + Z_1(x)$.
- MF:

$$\log\{p_2(x)/[1-p_2(x)]\} = d_2(x)^\top\beta_2 + S_2(x) + Z_2(x) + \gamma f(p_1(x))$$

  where $f : [0,1] \to I \subseteq \mathbb{R}$ continuous non-decreasing.
- Separate models if $\gamma = 0$.
- If $S_2(x) = 0$, for all x, and $f(p_1(x)) = \log\{p_1(x)/[1-p_1(x)]\}$, we recover Crainiceanu, Diggle and Rowlingson (2008).

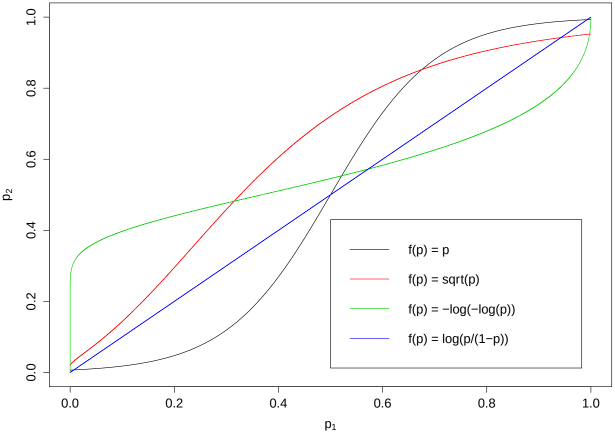
An empirical approach (1)

- ICT: $\log\{p_1(x)/[1-p_1(x)]\} = d_1(x)^\top\beta_1 + S_1(x) + Z_1(x)$.
- MF:

$$\begin{aligned}\log\{p_2(x)/[1-p_2(x)]\} &= d_2(x)^\top\beta_2 + S_2(x) + Z_2(x) + \\ &\quad \gamma f(p_1(x))\end{aligned}$$

  where $f:[0,1]\to I \subseteq \mathbb{R}$ continuous non-decreasing.
- Separate models if $\gamma = 0$.
- If $S_2(x) = 0$, for all x, and $f(p_1(x)) = \log\{p_1(x)/[1-p_1(x)]\}$, we recover Crainiceanu, Diggle and Rowlingson (2008).
- If $\gamma = 1$ and $f(p_1(x)) = \log\{p_1(x)/[1-p_1(x)]\} - d_1(x)^\top\beta_1 - Z_1(x)$, we recover Giorgi, Sesay, Terlouw and Diggle (2015).

An empirical approach (2)



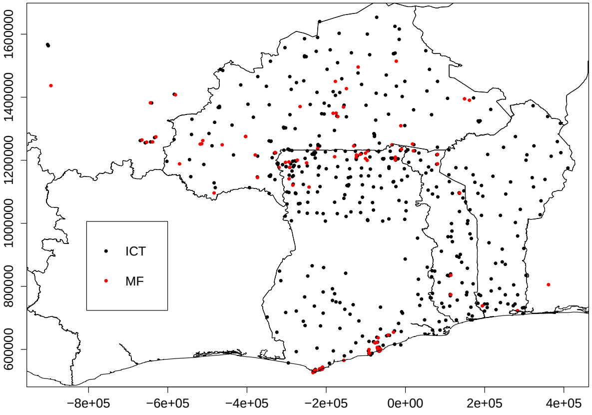
Application: LF mapping in West Africa

Application: LF mapping in West Africa

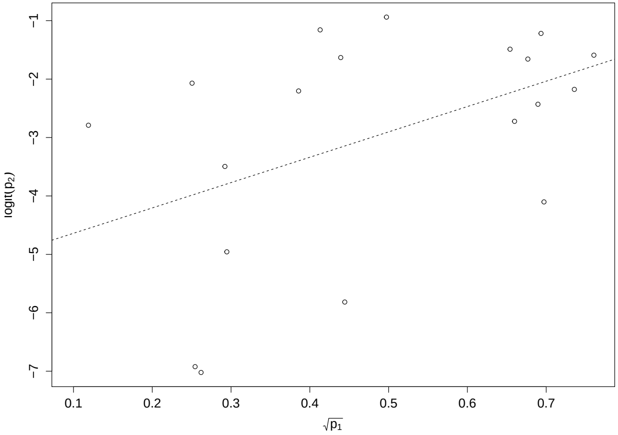
- MF and ICT surveys conducted from 1997 to 2003.

Application: LF mapping in West Africa

- MF and ICT surveys conducted from 1997 to 2003.
- 479 ICT surveys; on average 61 individuals sampled per village.
- 90 MF surveys; on average 245 individuals sampled per village.



Empirical relationship



Estimation (1)

Semi-mechanistic model with density-dependence

- $\log\{\lambda(x)\} = \mu + S(x) + Z(x)$, $r(x) = \alpha\lambda(x)^{\gamma}$ for all x.

Estimation (1)

Semi-mechanistic model with density-dependence

- $\log\{\lambda(x)\} = \mu + S(x) + Z(x)$, $r(x) = \alpha\lambda(x)^{\gamma}$ for all x.
- cov$\{S(x), S(x+h)\} = \sigma^2 \exp(-\|h\|/\phi)$,
  var$\{Z(x)\} = \nu^2\sigma^2$.

Estimation (1)

Semi-mechanistic model with density-dependence

- $\log\{\lambda(x)\} = \mu + S(x) + Z(x)$, $r(x) = \alpha\lambda(x)^{\gamma}$ for all x.
- cov$\{S(x), S(x+h)\} = \sigma^2 \exp(-\|h\|/\phi)$,
  var$\{Z(x)\} = \nu^2\sigma^2$.

Term	Estimate	95% CI
μ	-2.562	(-3.991, -1.132)
α	0.722	(0.636, 0.820)
γ	0.106	(0.010, 0.203)
σ^2	2.974	(1.392, 6.355)
ϕ	275.995	(117.374, 648.980)
ν^2	0.201	(0.091, 0.445)

Estimation (2)

Empirical model

- $\log\{p_1(x)/[1-p_1(x)]\} = \mu_1 + S_1(x) + Z_1(x),$
  $\log\{p_2(x)/[1-p_2(x)]\} = \mu_2 + S_2(x) + Z_2(x) + \gamma\sqrt{p_1(x)}$

Estimation (2)

Empirical model

- $\log\{p_1(x)/[1 - p_1(x)]\} = \mu_1 + S_1(x) + Z_1(x)$,
  $\log\{p_2(x)/[1 - p_2(x)]\} = \mu_2 + S_2(x) + Z_2(x) + \gamma\sqrt{p_1(x)}$
- cov$\{S_i(x), S_i(x+h)\} = \sigma_i^2 \exp(-\|h\|/\phi_i)$,
  var$\{Z_i(x)\} = \nu_i^2\sigma_i^2$, i=1,2.

Estimation (2)

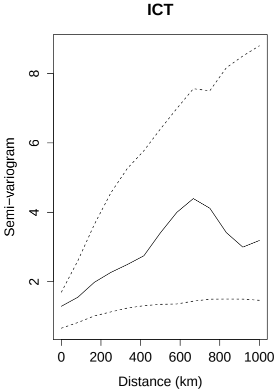
Empirical model

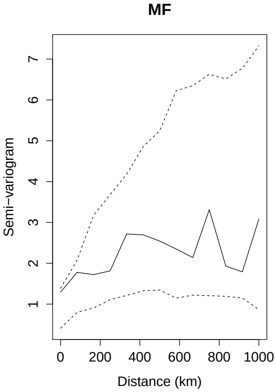
- $\log\{p_1(x)/[1-p_1(x)]\} = \mu_1 + S_1(x) + Z_1(x)$,
  $\log\{p_2(x)/[1-p_2(x)]\} = \mu_2 + S_2(x) + Z_2(x) + \gamma\sqrt{p_1(x)}$
- cov$\{S_i(x), S_i(x+h)\} = \sigma_i^2 \exp(-\|h\|/\phi_i)$,
  var$\{Z_i(x)\} = \nu_i^2\sigma_i^2$, i=1,2.

Term	Estimate	95% CI
μ_1	-2.244	(-4.201, -0.287)
μ_2	-3.055	(-3.763, -2.348)
γ	0.476	(0.200, 0.752)
σ_1^2	4.205	(1.667, 10.608)
ϕ_1	354.055	(126.608, 990.106)
ν_1^2	0.172	(0.066, 0.449)
σ_2^2	1.796	(0.909, 3.550)
ϕ_2	82.555	(34.412, 198.052)
ν_2^2	0.228	(0.069, 0.760)

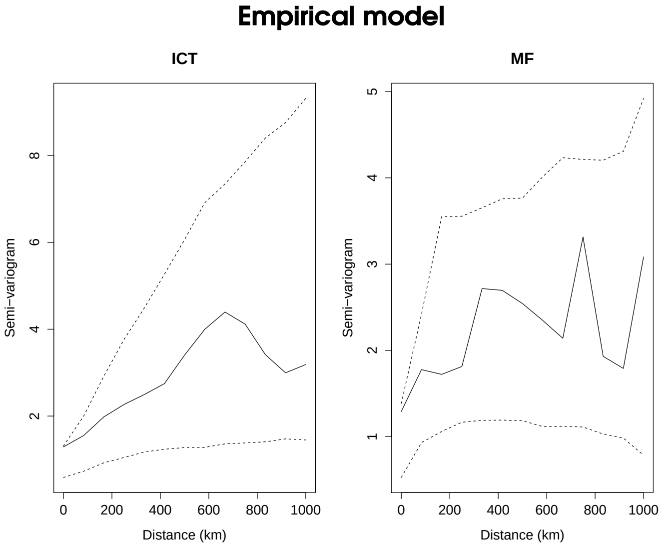
Model diagnostic (1)

Semi-mechanistic model with density-dependence



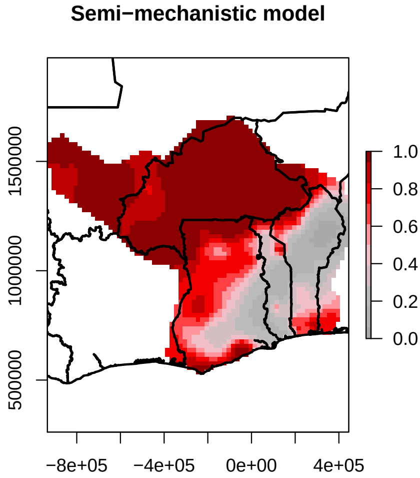


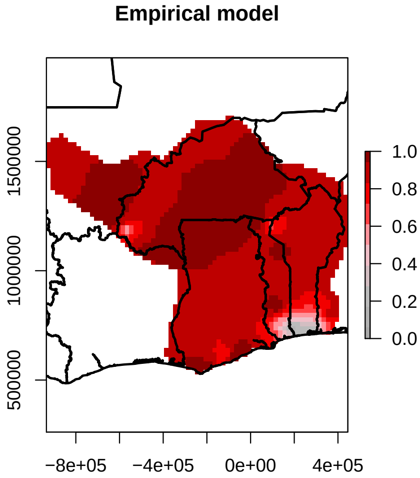
Model diagnostic (2)



Exceeding 1% MF prevalence

Exceeding 1% MF prevalence





Discussion

- Which model is the best with respect to the scientific knowledge?

Discussion

- Which model is the best with respect to the scientific knowledge?
- **Simulation study:** empirical model provides robust inferences against the misspecification of f.
- **Simulation study:** misspecification of the model may still yield accurate point predictions but actual coverage of CI may be very different from the nominal.

Discussion

- Which model is the best with respect to the scientific knowledge?
- **Simulation study:** empirical model provides robust inferences against the misspecification of f.
- **Simulation study:** misspecification of the model may still yield accurate point predictions but actual coverage of CI may be very different from the nominal.

Thank you for your attention!

Bibliography

1. M. A. Irvine, S. M. Njenga, S. Gunawardena, C. N. Wamae, J. Cano, S. J. Brooker, and T. D. Hollingsworth. **Understanding the relationship between prevalence of microfilariae and antigenaemia using a model of lymphatic filariasis infection**. Trans R Soc Trop Med Hyg (2016) 110(5): 317 doi:10.1093/trstmh/trw024
2. C. Crainiceanu, P.J. Diggle, and B.S. Rowlingson. **Bivariate modelling and prediction of spatial variation in Loa loa prevalence in tropical Africa (with Discussion).** (2008) Journal of the American Statistical Association, 103, 21-43.
3. E. Giorgi, S.S. Sesay, D.J. Terlouw and P.J., Diggle. **Combining data from multiple spatially referenced prevalence surveys using generalized linear geostatistical models.** (2015) Journal of the Royal Statistical Society A 178, 445-464.