Semi-mechanistic model and its implementation in R *

Claudio Fronterrè *London School of Hygiene and Tropical Medicine*

This document provides details for the semi-mechanistic model developed to estimate MF prevalence when two different diagnostic tools are used. It also shows how to implement it in R.

Keywords: lymphatic filariasis, antigeamia, geostatistics

1 Introduction

Lymphatic filariasis (LF) is a mosquito-borne neglected tropical disease targeted for global elimination by 2020. The majority of global cases are caused by three species of nematode worms: *Wuchereria bancrofti*, *Brugia malayi* and *Brugia timori*. These filariae parasites are transmitted by various species of mosquito vectors from the genera Anopheles, Aedes, Culex, Mansonia and Ochlerotatus. In recent years, the mapping of LF has been greatly facilitated by the use of simple and rapid detection tests for *W. bancrofti* (antigen-based test) and *Brugia* (antibody-based test), based on the immuno-chromatographic test (ICT card test), which avoids the need to collect blood at night and the time-consuming preparation and examination of blood slides. While it is known that estimates of antigenaemia are generally higher than estimates of microfilaraemia (MF), the extent and spatial heterogeneity of this relationship is not clear. Even if the scientific output of interest is the prevalence of microfilaraemia, the small number of mapping surveys that measure mf is low and it is decreasing due to the diffusion and cost-effectivness of ICT tests. Our goal is to use the abundace of ICT prevalence surveys and the relationship between ICT and MF prevalence to predict microfilaraemia prevalence at unobserved locations.

2 Modeling framework

We can define our set of data as follow:

$$\mathcal{D} = \left\{ \left(x_{i,j}, n_{i,j}, y_{i,j}\right) : x_{i,j} \in A \right\}, i = 1, \ldots, n_j, j = 1, 2;$$

where x is the geograpich location of the mapping survey, y is the number of people infected out of n examined and j indicate which type of prevalence was measured, $1 = MF$ and $2 = ICT$. Our final target of estimation is the predictive distribution of MF prevalance given ICT prevalence $[Y_1 \mid Y_2, S]$. We can consider the observed prevalence as the realisation of a binomial random variable

$$Y_{i,j} \mid S_j(x_i j), Z_{i,j} \sim \text{Binomial}(n_i, j, p_j(x_{i,j})).$$

We use biological information to define the probability of being tested as postive

$$p_j(x) = \begin{cases} 1 - \exp\left\{-\lambda(x)\left[1 - \exp(-\alpha)\right]\right\} & j = 1 \\ \phi\left\{1 - \exp\left[-\lambda(x)\right]\right\} & j = 2 \end{cases}$$

where α is a parameter that control the reproductive rate of MF, ϕ is the sensitivy of the ICT test and $\lambda(x)$ is the mean number of adult worms in a sampled individual. The last one is assumed to

*Corresponding author: claudio.fronterre@lshtm.ac.uk.

vary spatially using the following specification

$$\log \lambda(x) = d(x)^{\mathrm{T}}\beta + S(x) + Z.$$

2.1 Lieklihood

Let θ be the vector of model parameters to be estimated. The log-likelihood for this model is

$$l(\theta) = \sum_{j=1}^{2} \left[n_j \sum_{i=1}^{n} \log\left(1 - p_{i,j}\right) + \sum_{i=1}^{n} \log\left(\frac{p_{i,j}}{1 - p_{i,j}}\right) \right].$$

3 Implementation in R