









# Data-driven modelling and spatial complexity supports heterogeneity-based integrative management for eliminating *Simulium neavei*-transmitted river blindness

Edwin Michael , Morgan E. Smith , Brajendra K. Singh , Moses N. Katarbarwa , Edson Byamukama , Peace Habomugisha , Thomson Lakwo , Edridah Tukahebwa  & Frank O. Richards

Concern is emerging regarding the challenges posed by spatial complexity for modelling and managing the area-wide elimination of parasitic infections. While this has led to calls for applying heterogeneity-based approaches for addressing this complexity, questions related to spatial scale, the discovery of locally-relevant models, and its interaction with options for interrupting parasite transmission remain to be resolved. We used a data-driven modelling framework applied to infection data gathered

*Simulium neavei*-transmitted onchocerciasis, a macroparasitic disease that causes river blindness in Western Uganda and other regions of Africa. We demonstrate that our Bayesian-based data-model assimilation technique is able to discover

such as infection breakpoints and required durations of drug interventions for achieving elimination

operate at the larger focus level, although intriguingly including vector control overcame this variability. These results show that data-driven modelling based on spatial datasets and model-data fusing methodologies will be critical to identifying both the scale-dependent models and heterogeneity-based options required for supporting the successful elimination of *S. neavei*-borne onchocerciasis.

In recent years, there has been growing appreciation of the role that mathematical models can play in guiding the control or eradication of the major preventable helminthic diseases, ranging from soil-transmitted helminthiasis to schistosomiasis, lymphatic filariasis and onchocerciasis. This recognition has been catalysed by the expectation that quantitative mathematical models can provide a robust scientific basis for predicting the course of infection result in a more effective and targeted control strategy.

4,13–15





Kolmogorov-Smirnov tests comparing posterior parameter distributions between site-specific and focus ABR models, however, showed that the distributions for these parameters are significantly different, indicating their sensitivity to the choice of ABR type (whether focus or site-specific) used in the present simulations (Table 3). Interestingly, most of the parameters that are affected by this sensitivity to the specific ABR used are those related to exposure ( $H_b$ ,  $\beta_1$ ,  $\beta_2$ ,  $H_{lin}$ ), suggesting that these ecological parameters can compensate for the change in ABR values employed in order to reproduce the observed mf infection profiles. Note the fact that such different parameter and ABR combinations are able to reproduce the observed age infection profiles in a site equally well indicates that complex interactions and trade-offs exist between parameters or that multiple model representations<sup>20,43,44</sup> can act to provide acceptable simulations for the onchocerciasis transmission system. This further emphasizes the











dynamics within an endemic zone suggests that management of *S. neavei*-transmitted onchocerciasis elimination will need to be sensitive to variable parasite transmission and extinction dynamics at the focus level rather than be based on a paradigm of uniformity that ignores this spatial heterogeneity.

The above conclusions are clearly conditioned on the ability of our modelling framework to reliably capture the dominant transmission dynamics of *S. neavei*-transmitted onchocerciasis in an endemic focus. While our BM-based modelling algorithm combines the advantageous features of mechanistic and statistical approaches to improve the estimation of local models for facilitating forecasts of interventions applied under a variety of field conditions, it is dependent, as for any data-driven predictive system, on the model structure employed, estimation procedure, and on the data used for facilitating model discovery<sup>16,17</sup>. Although our BM framework primarily focused attention on addressing parameter uncertainty with data, we note here firstly that the present model is based on previously established population models of onchocerciasis transmission<sup>7,47</sup>, with appropriate structural extensions made with regard to population-averaged mf uptake and larval development in the *S. neavei* vector host as well as the operation of different forms of host immunity in populations<sup>48</sup>. Furthermore, we have also secondly included all previously suggested density-dependent functions that are thought to govern onchocerciasis transmission but do not make any *a priori* assumptions concerning their occurrence using data, instead, to determine the operation of these functions in a site, which allows for a degree of updating for model structures applicable to a particular setting.

By contrast to conventional model calibration practices (which principally focus on adjustment of parameters until discrepancy between model outputs and observed data is minimized), the Bayesian basis of our Monte-Carlo BM-based model estimation technique, as we have demonstrated previously<sup>25,30,31,49</sup>, is also more informative given that it not only updates our knowledge of model parameters (and structure) for a site but also facilitates predictions along with associated uncertainties for modelled output variables<sup>38,50</sup>. In addition, it is also vitally premised on the expectation that obtaining a single, optimal, set of acceptable parameters for describing a complex multi-parameter ecological system is problematic, as parameter equifinality means that often several







elimination of complex vector-borne macroparasitic diseases, such as *S. neavei*-transmitted onchocerciasis. We also show that modelling frameworks that couple spatial datasets with data-model assimilation methods will be required if the appropriate scale-dependent local models are to be discovered and used to facilitate the making of predictions of the impact of implemented or proposed interventions that take a fuller account of spatial com-

assumption that the density of infective stage larvae in the vector population reaches a dynamic equilibrium rap-

and input drivers to constrain model parameters and quantify model error<sup>30,82</sup>, thereby improving the discovery of behaviourally acceptable complex dynamical models applicable to a given setting<sup>20</sup>. Here, we build on one such strategy that facilitates the integration of field observations on onchocerciasis baseline mf prevalence and ABR in sentinel communities with simulation model outputs to estimate the values (and uncertainty) in the model parameters applicable for reproducing local parasite transmission dynamics. We then apply the locally calibrated model to undertake: (1) quantification of the infection endpoints or elimination thresholds applicable in each setting; and (2) based on model predicted mf prevalence trajectories, examine the effectiveness of various

exceedance calculations to quantify the values of mf breakpoint prevalence thresholds reflecting a 95% elimination probability for use as the infection elimination target in this work. We used both the observed focus-level ABRs well as the estimated site-specific ABRs as required (see below) to estimate the site-specific worm/mf breakpoint and TBR values used in this analysis.

Previous studies have shown that the key environmental driver that governs pattern-process relationships in the case of vector-borne macroparasitic infections, and hence that primarily constrains a model's parameter space so that the effects of local transmission conditions are captured, is the vector biting rate obtaining in a community<sup>4,7,25,30</sup>. However, in this study, the baseline data for each study village included information on this driving variable only at the higher aggregate focus level as described above. Although the use of such aggregated ABR to obtain the best-fit models within each sentinel village in a focus may offer a second-best option in the case of lack of information on vector abundance at the local level, it is apparent that this approach could introduce significant bias if the sentinel villages differed markedly in their local ABR values. We examined this possibility by using the models fitted to the mf prevalence data in each site in order to 1) estimate the corresponding site-specific ABRs, and 2) following this, undertaking a comparison of the values obtained against each measured focus-level ABR. To perform this exercise, each of the randomly sampled parameter vectors was used to obtain plausible ABR values by running the model under a standard root-finding algorithm for which the model-generated overall mf prevalence matched the observed baseline community mf prevalence within a tolerance limit of 0.001 or lower. This root-finding algorithm was implemented in Matlab using its built-in *fzero* function<sup>4</sup>. The same procedure was also used to estimate the missing ABR data for the Bwindi sites.

The impact of MDA was modelled by assuming that anti-malarial treatment with the currently used ivermectin drug regimens acts by killing certain fractions of the populations of adult worms and mf instantly after drug administration, as well as reducing the fertility of surviving female adult worms<sup>86–88</sup>.







