Published in final edited form as:

Adv Parasitol. 2011 ; 74: 267–296. doi:10.1016/B978-0-12-385897-9.00005-7.

The Applications of Model-Based Geostatistics in Helminth Epidemiology and Control

Ricardo J. Soares Magalhães^{*}, Archie C.A. Clements^{*,†}, Anand P. Patil[‡], Peter W. Gething[‡], and Simon Brooker^{§,¶}

^{*}University of Queensland, School of Population Health, Herston, Queensland, Australia

[†]Australian Centre for International and Tropical Health, Queensland Institute of Medical Research, Herston, Queensland, Australia

[‡]Department of Zoology, University of Oxford, Oxford, United Kingdom

[§]Kenya Medical Research Institute-Wellcome Trust Research Programme, Nairobi, Kenya

[¶]London School of Hygiene and Tropical Medicine, Department of Infectious and Tropical Diseases, London, United Kingdom

Abstract

Funding agencies are dedicating substantial resources to tackle helminth infections. Reliable maps of the distribution of helminth infection can assist these efforts by targeting control resources to areas of greatest need. The ability to define the distribution of infection at regional, national and subnational levels has been enhanced greatly by the increased availability of good quality survey data and the use of model-based geostatistics (MBG), enabling spatial prediction in unsampled locations. A major advantage of MBG risk mapping approaches is that they provide a flexible statistical platform for handling and representing different sources of uncertainty, providing plausible and robust information on the spatial distribution of infections to inform the design and implementation of control programmes. Focussing on schistosomiasis and soil-transmitted helminthiasis, with additional examples for lymphatic filariasis and onchocerciasis, we review the progress made to date with the application of MBG tools in large-scale, real-world control programmes and propose a general framework for their application to inform integrative spatial planning of helminth disease control programmes.

5.1. INTRODUCTION

Effective control of human helminth infections requires reliable estimates of the geographical distribution of infection and the size of populations requiring intervention (Boatin and Richards, 2006; Brooker and Michael, 2000; Brooker et al., 2006b; Molyneux, 2009). For the purposes of control planning, nationwide surveillance data are desirable, but few endemic countries have suitably detailed data (Brooker et al., 2000b). To address this paucity of data, research over the past decade has explored ways to maximise the usefulness of available data based on disease mapping and prediction (Brooker, 2002, 2007; Brooker and Michael, 2000; Brooker et al., 2006b,c; Simoonga et al., 2009). Most recently, these predictive approaches have employed Bayesian model-based geostatistics (MBG) which embeds classical geostatistics in a generalised linear modelling framework. Using this approach, relationships and associated uncertainty between infection outcomes and covariates are estimated and the resultant model is used to predict the outcome at unsampled locations (Diggle, Tawn et al., 1998). This approach has the advantage over traditional spatial prediction methods of providing a robust and comprehensive handling of spatial structure and the uncertainty associated with predicted infection patterns.

This review focuses on human helminth infections: schistosomiasis, intestinal nematodes (or soil-transmitted helminths, STH), lymphatic filariasis (LF) and onchocerciasis; but it is important to recognise the increasing number of applications of MBG to spatial modelling of malaria infection (Craig et al., 2007; Diggle et al., 2002; Gosoni et al., 2006, 2009; Hay et al., 2009; Kazembe et al., 2006; Noor et al., 2008, 2009; Raso et al., 2009b; Silue et al., 2008), malaria-related mortality (Gemperli et al., 2004) and malaria entomological inoculation rates (Gemperli et al., 2006a,b).

The primary aim of this review is to demonstrate the applications of MBG to helminth epidemiology and encourage its wider application in helminth disease control programmes. The first section highlights the disease burden of helminth infections in SSA and identifies the main treatment strategies. The second section examines the importance of mapping in guiding helminth control. The third section introduces the principal concepts that underpin MBG. This is followed by a description of the survey data requirements for MBG, before showing how survey and satellite-derived environmental data have been integrated into an MBG platform to establish and predict species-specific prevalence and intensity distributions, and describes how these tools could be extended to accommodate sampling uncertainty and greater biological realism. Finally, we review how these tools have already helped inform large-scale control programmes and look forward to their future potential application. The search strategy and selection criteria of the review are shown in Box 5.1.

5.2. DISEASE BURDEN AND INTERVENTION STRATEGIES

Helminths are some of the most common infections of humans. In sub-Saharan Africa (SSA), 740 million individuals are estimated to be infected with soil-transmitted helminths (*Ascaris lumbricoides*, *Trichuris trichiura*, and the hookworms *Necator Americanus* and *Ancylostoma duodenale*) (de Silva et al., 2003), 207 million with schistosomiasis (*Schistosoma haematobium* and *S. mansoni*) (Steinmann et al., 2006), 50 million with LF due to *Wuchereria bancrofti* (Michael and Bundy, 1997), and 18–37 million with onchocerciasis due to *Onchocerca volvulus* (Basanez et al., 2006). All of these parasites can be effectively treated with single dose oral therapies that are safe, inexpensive and required at periodic intervals. STH infections are treated with albendazole or mebendazole (Gulani et al., 2007; Keiser and Utzinger, 2008; Taylor-Robinson et al., 2007), whilst

BOX 5.1

Search strategy and selection criteria

Data for this review were obtained from publications identified by a systematic search of PubMed, focusing on those published from 2001 to 2009. Search terms for each parasite included:

- Schistosomiasis: (schistosoma or schistosomiasis or bilharziose) and Africa and spatial
- Onchocerciasis: (onchocerca or onchocerciasis) and Africa and spatial
- Trichuriasis: (trichuris or trichuriasis or tricuriase) and Africa and spatial
- Ascariasis: (ascaris or ascariasis or ascariase) and Africa and spatial
- Lymphatic filariasis: (lymphatic or bancroftian) and filariasis and Africa and spatial
- Hookworm: (hookworm or Necator or Ancylostoma) and Africa and spatial

Abstracts of English, French, Portuguese and Spanish language papers were read and considered for inclusion, although only English language papers were selected for the final review. Secondary, manual searches of the cited references of these articles were conducted and relevant articles were included.

schistosomiasis is treated with praziquantel (Richter, 2003). Lymphatic filariasis is treated using albendazole with diethylcarbamazine or ivermectin, and ivermectin is the choice of drug for onchocerciasis (Olsen, 2007). Treatment is typically implemented through mass chemotherapy whereby the entire at-risk population is treated, as part of either school or community-based campaigns.

A number of international initiatives have supported mass school-based treatment for STH infection and schistosomiasis, including Deworm the World (www.dewormtheworld.org) and Children Without Worms (www.childrenwithoutworms.org) for STH infection, and the Schistosomiasis Control Initiative (SCI; www.sci-ntds.org), initially for schistosomiasis and STH (Fenwick et al., 2009). Global control of filariasis is coordinated by the Global Programme to Eliminate Lymphatic Filariasis, a public–private partnership led by WHO, and which has provided treatment with ivermectin and albendazole to more than 1900 million people in 48 countries worldwide (Hooper et al., 2009). The control of onchocerciasis in SSA is overseen by the African Programme for Onchocerciasis Control (APOC; Boatin and Richards, 2006; Boatin et al., 1998). This programme has to date treated 55 million people with ivermectin in 16 participating countries.

Several defined measures of helminth transmission are valuable to guide the implementation of the control programmes described above. The most commonly measured is the prevalence of infection (the proportion of individuals infected). A second key measure is the intensity of infection (the worm burden) which is estimated based on quantitative egg counts or blood smears. The relative ease in collecting prevalence data means that the decision on where to implement control is typically based on whether the prevalence of infection exceeds some species-specific threshold. For STH and schistosomiasis, where the goal of treatment is morbidity control, mass treatment has been recommended where the prevalence of infection exceeds 20% among school children (WHO, 2002, 2006). Regarding LF, for which the goal is elimination, the threshold is prevalence >1%, whilst mass treatment with ivermectin is implemented in areas where prevalence of onchocerciasis is >20% (WHO, 2002). Regardless of the treatment threshold, the implementation of helminth control requires evidence-based maps of infection prevalence.

5.3. THE ROLE OF MAPPING IN HELMINTHOLOGY

The inherent spatial heterogeneity of infection varies between individual helminth species. Generally, the more complex the life cycle, the more spatially heterogeneous infection patterns appear. For example, in East Africa, schistosomiasis, LF or onchocerciasis, for which transmission involves either an intermediate host or vector, typically have a focal distribution, whereas STH are more widely distributed in space owing to their direct transmission life cycle (Brooker, 2007; Brooker et al., 2004; Gyapong et al., 2002; Sturrock et al., 2009).

To help reduce the costs of prevalence surveys, effort has been invested in developing rapid assessment methods to determine the prevalence of infection as inexpensively and as quickly as possible. For example, to identify communities at high risk of onchocerciasis, requiring mass treatment with ivermectin, APOC implements rapid epidemiological mapping of onchocerciasis (REMO; Noma et al., 2002). This technique provides data on the distribution and prevalence of onchocerciasis, enabling delineation of zones of varying

endemicity. For other diseases, similar approaches have been developed, including the rapid geographical assessment of bancroftian filariasis (RAGFIL) method (Gyapong and Remme, 2001; Gyapong et al., 2002; Srividya et al., 2002) and rapid assessment procedure for loiasis (RAPLOA; Takougang et al., 2002; Thomson et al., 2004). Other rapid mapping tools include school-based blood in urine questionnaire surveys (Clements et al., 2008a,b; Lengeler et al., 2002) and parasitological surveys based on lot quality assurance sampling (LQAS; Brooker et al., 2005, 2009). For a review of rapid mapping techniques and tools the reader is referred to Brooker et al. (2009).

To augment approaches to rapid mapping and also address the absence of suitable data in many settings, spatial prediction methods, based on statistical relationships between individual and environmental predictors and observed risk of infection, are increasingly being used. Advances in geographical information systems (GIS) and remote sensing (RS) technologies over the past 20 years have greatly facilitated the explanation and prediction of patterns of parasitic disease risk by providing a platform for integration of survey data with data on environmental and socioeconomic determinants (Hay et al., 2006; Robinson, 2000). Data warehousing and expansion of the internet have made many datasets on potential environmental and socio-economic predictor variables more accessible, with a wide range of datasets now freely available from on-line sources (e.g. <http://srtm.csi.cgiar.org/>; <http://www.worldclim.org/>; <http://sedac.ciesin.columbia.edu/gpw/>; <http://www.fao.org/geonetwork/srv/en/main.home>). Typically for spatial analysis, field survey data, which contain information on either prevalence or intensity of infection and individual-level covariates such as age and sex, are assembled in a GIS and linked to community or school-level RS environmental and socio-economic predictor data. The linked dataset can then be exported from the GIS for multivariable modelling, with particular recent attention being paid to MBG (Diggle et al., 1998). For a review on non-Bayesian approaches to helminth mapping, the reader is referred to reviews by Brooker et al. (2006b) and Simoonga et al. (2009).

5.4. PRINCIPLES OF MODEL-BASED GEOSTATISTICS

A central feature of MBG is that it can take into account spatial dependence, also known as spatial autocorrelation (Box 5.2). This is the phenomenon that values at nearby locations tend to be more similar than those further apart (Tobler, 1970). Standard regression techniques rely on an assumption of conditional independence in model residuals. When handling spatially autocorrelated data, this assumption is often violated, with model residuals likely to display some degree of spatial autocorrelation (Kuhn, 2007), presenting a particular problem for spatial risk mapping (Dormann, 2007).

An established approach for handling spatially autocorrelated data stems from classical geostatistics, which uses kriging for spatial prediction (Cressie, 1990; Wackernagel, 2003). This is a group of techniques which allows smoothing of values observed at sampled locations and prediction at unsampled locations. This method of interpolation uses a semi-variogram (see Box 5.2) to define the spatial variation of the data and minimise

BOX 5.2

The nomenclature of spatial dependence

One way of graphically describing the extent of spatial dependence in point-referenced data is via the semi-variogram. A semi-variogram is a mathematical function which describes the variability of a measure with location, by examining the variation in observations with distance between all pairs of sampled locations. The semi-variogram is described by at least three parameters; the nugget, partial sill and range (Cressie, 1993).

The nugget represents spatially random (i.e. uncorrelated) variation which could arise due to natural random variation, very small-scale spatial variability and/or measurement error. The partial sill represents spatially autocorrelated variation which could arise due to spatial heterogeneity in important, unmeasured drivers of transmission (i.e. factors not included as covariates in the model and/or the requirement for close spatial proximity between infectious and susceptible individuals for transmission events to occur (manifested by disease clusters)). The range is the separating distance at which spatial correlation ceases to occur and is an indication of the size of disease clusters.

the error variance associated with predicted values (Cressie, 1990). Classical geostatistics is best suited to Gaussian (i.e. normally distributed) outcomes and can encounter difficulties in quantifying prediction uncertainty for non-Gaussian outcomes such as proportions (e.g. prevalence of infection) or counts (e.g. number of eggs per gram of faeces). It was this limitation that MBG was primarily developed to overcome (Diggle et al., 1998). Additionally, MBG is generally implemented in a Bayesian inferential framework, thereby providing an intuitive interpretation of parameter uncertainty whilst explicitly modelling spatial autocorrelation and, most importantly, allowing a formal expression of uncertainty in the prediction estimates (Diggle et al., 1998). The application of Markov chain Monte Carlo simulation (MCMC) for model fitting helps to address the considerable computational challenges previously incurred when computing the high-dimensional integrals necessary for Bayesian analysis.

The outputs of Bayesian modelling are probability distributions, termed predictive posterior distributions, which represent the probability of a parameter of interest taking values from within a plausible range. This inferential framework has important practical implications in risk mapping because posterior predictive distributions can be derived for both the parameters (which include the spatial autocorrelation parameters and the coefficients of covariates) and the epidemiological outcome of interest (e.g. prevalence or intensity of infection) at unsampled locations, which classical geostatistics can achieve only in special circumstances (Lawson, 2009). Posterior predictive distributions for the infection outcome of interest can be computed on a pixel-by-pixel basis, providing either posterior predictive distributions that are independent of neighbouring values or jointly. Sampling from the joint posterior distribution incurs very substantial computational expense (Lawson, 2009), but has the important advantage of allowing spatial aggregation of predictions, such as the mean or sum of the predicted values of the infection measure of interest over a spatial region (Gething et al., 2010).

The incorporation of uncertainty into the modelling framework and the expression of uncertainty of predictions are particular strengths of MBG. Uncertainty is an intrinsic feature of all spatial predictions at unsampled locations based on data observed at sampled locations and has multiple sources, including sampling error, measurement error of both outcomes or covariates, as well as prediction errors at unsampled locations (Agumya and Hunter, 2002; Leonardo et al., 2008). Bayesian methods are ideally suited to dealing with such multiple sources of uncertainty and also permit incorporation of additional sources of information (e.g. prior knowledge about natural history of infection). Uncertainty in spatial prediction is typically explored by examining the posterior distributions: those with large variances are indicative of lower predictive precision and higher associated uncertainty. Typically, precision tends to be lower in areas where there are less data or in areas where the data themselves are highly heterogeneous over short distances.

The formal representation of uncertainty afforded by MBG models has practical use for control programmes. Different percentiles of the posterior distribution (upper and lower

quartiles or 95% credible intervals) can be mapped, thereby demonstrating the range of plausible values for each location. The flexibility afforded by MGB also allows demonstration of the probability that predicted prevalence is above (or below) a given mass treatment threshold by constructing probability contour maps (see Section 5.2).

Box 5.3 presents the main steps in implementing MBG for the prediction of helminth distributions. In the MBG framework, spatial variation is said to occur over two scales: large-scale variation (so-called *first order* variation, or trend); and local spatial variation (so-called *second order* variation). First order variation can be modelled using individual covariates (such as age, sex, or anthropometric variables) or spatially contextual covariates (e.g. climate, proximity to water bodies), while *second order* variation is modelled by introducing location-specific random effects, structured as a multivariate normal-distributed random field with a correlation matrix defined by a spatially decaying autocorrelation function.

As will be seen in subsequent sections, MBG can incorporate different epidemiological information (inputs) and produce a range of prediction

BOX 5.3

Example of general steps for geostatistical modelling of helminth infections

Step 1

Generally an initial candidate set of individual-level and environmental/climatic covariates is considered for inclusion in the models. Individual covariates could include age and sex recorded during field surveys. The value of the environmental/climatic covariates for each survey location is extracted in a GIS. Variable selection is made using fixed-effects univariable logistic regression models in a standard statistical package with backwards elimination. All variables which have a Wald's $P > 0.2$ are selected for inclusion in a final model.

Step 2

The residuals of the final model are examined for spatial autocorrelation using semi-variograms (Box 5.2.) in R version 2.4.0 (R Development Core Team). When spatial autocorrelation is apparent, this means that models incorporating a spatial dependence component (i.e. a geostatistical random effect) should be most appropriate.

Step 3

Spatial models of prevalence of infection, intensity of infection and co-infection can be developed in WinBUGS version 1.4 (MRC Biostatistics Unit, Cambridge, and Imperial College London, UK). The prevalence models shown here were logistic regression models and a geostatistical random effect that modelled spatial correlation using an isotropic, stationary exponential decay function (Diggle et al., 1998).

Step 4

Validation can be performed by dividing the original survey locations into four random subsets and sequentially withholding the data from one subset (the validation subset) while building the models with the remaining data (the training subset) and predicting prevalence of infection for the validation locations. Model calibration and discrimination can be determined by using the area under the curve of the receiver operating characteristic. Final model predictions (mean prevalence, upper and lower 95% credible intervals) are mapped in the GIS.

Steps 5 and 6

In our case, prediction of the spatial distribution of helminth infections was based on multivariable modelling at the nodes of a 0.1×0.1 decimal degrees ($\sim 12 \times 12$ km) grid covering the study areas (prediction locations). This can be done in WinBUGS using the spatial.unpred command, which implements an interpolation function (kriging), in our case for the spatial random effects; this function allows prediction without considering predicted values at neighbouring locations (marginal prediction). Predicted prevalence of infection is generally calculated by adding the interpolated random effect to the sum of the products of the coefficients for the covariates and the values of the covariates at the prediction locations. The overall sum was then back-transformed from the logit scale to the prevalence scale, giving prediction surfaces for prevalence of each type of helminth infection in each age and sex group. For an initial assessment of parameter convergence, the initial set of iterations (in our case the first 4000 iterations) is not considered (burn-in period). This is followed by another set of iterations until convergence (in our case 1000 iterations) where values for the intercept and coefficients were stored. Diagnostic tests for convergence of the stored estimates of parameter values are undertaken by visual examination of history and density plots of the model runs or chains. Once convergence is successfully achieved (in our case after 5000 iterations) the model was run for a further 10,000 iterations, during which predicted prevalence at the prediction locations was stored for each age and sex group. Inference is made by assessing posterior distributions of model parameters in terms of the posterior mean and 95% credible interval, which represents the values within which the true value occurs with a probability of 95%.

maps (outputs), and so provide a coherent planning framework. Figure 5.1 presents a potential framework for the use of MBG tools in the design and implementation of control

programmes targeting human helminth infections. In some cases, control managers may not be interested in integrating the disease control programme with other diseases but rather use an MGB application that allows single disease risk mapping, assessment of the geographic variation of disease risk and estimation of resource needs for a single-disease control programme—this can be achieved by prevalence mapping. This is the most common approach to predictive mapping documented in the literature; one important planning advantage is that it allows enumeration of resource needs by combining data on population at risk. The main disadvantage is its limited use as an evaluation tool since prevalence is not the most appropriate indicator of changes in disease morbidity. Alternatively, mapping prevalence of intensity profiles, intensity of infection or clinical morbidity profiles can provide suitable indicators of infection morbidity levels and therefore has the added benefit of potentially being used as a control programme evaluation tool (See section 5.6.2).

5.5. DATA REQUIREMENTS FOR MBG

Any model is only as reliable as the data on which it is based. In turn, the most appropriate sampling design for data collection will depend on the intended purpose of the mapping exercise, which is linked to the objectives of the control programme. However, risk mapping is rarely based on data explicitly collected for the purpose of spatial prediction. Instead, data have often been collected for purposes other than spatial analysis, potentially limiting their usefulness for spatial prediction. Such problems might include inadequate sample size for estimating prediction model parameters, uneven spatial sampling density or incomplete coverage of the geographical area of interest, leading to low-precision predictions in some areas. Other challenges relate to difficulties in geo-locating sampling locations, necessitating retrospective geo-location using external, secondary data sources, which can introduce additional error.

Survey designs for risk mapping can take either a design-based or a model-based approach. In design-based sampling, the configuration of the sampling locations is random and the values at given locations are assumed to be fixed, whereas in model-based sampling the configuration of the sampling locations is fixed and the values at given locations are assumed to be realisations of a random variable (Brus and Gruijter, 1993, 1997). An example of data that were collected explicitly for helminth mapping using a design-based approach is the data obtained from national cross-sectional surveys supported by SCI in six African countries (Burkina Faso, Burundi, Ghana, Mali, Niger and Rwanda) and subnational data in Tanzania and Zambia (Fenwick et al., 2009). These surveys were conducted using standardised protocols and included a stratified cluster random sampling design (i.e. a design-based approach using probability-based sampling; Fig. 5.2). Specifically, schools or communities were selected from a sampling frame (obtained by government lists of schools or communities), and individuals were sampled within the selected units from an assembly of all individuals in a central location (typically, the school). This approach to spatial sampling sought to obtain a representative sample but also an adequate geographical coverage of the survey area.

A model-based approach to sampling takes into account the overall spatial variability of outcomes and measures of association with covariates; an MBG framework can be used to derive an optimal spatial design for prediction at unknown locations and for estimation of the spatial dependence (variogram) parameters while making appropriate allowance for parameter uncertainty (Diggle and Lophaven, 2006; Diggle et al., 1998). The main advantage of the model-based approach over a design-based approach is that the former can be used to derive an optimal spatial design by determining the number, dimensions and spatial arrangement of the sites that optimise the available data (Waller, 2002). The resulting design is typically a combination of a regular grid with additional points at shorter distances

to inform estimation of the spatial dependence parameters. The reader is referred to Diggle et al. (1998) and Diggle and Lophaven (2006) for further explanation of MBG approaches to survey design. Spatially explicit survey design is clearly an area that deserves more critical evaluation in helminth epidemiology and control.

5.6. MODEL-BASED GEOSTATISTICS APPLICATIONS IN HELMINTHOLOGY

5.6.1. Mapping prevalence of infection

MBG has been applied to the mapping of helminth infection at various spatial scales. Applications at national and subnational levels include: *S. mansoni* in western Cote d'Ivoire (Beck-Worner et al., 2007; Raso et al., 2005), Mali (Clements et al., 2009) and Tanzania (Clements et al., 2006a); *S. haematobium* in Mali (Clements et al., 2009) and Tanzania (Clements et al., 2006a); STHs in western Cote d'Ivoire (Raso et al., 2006a); and *Loa loa* infection in Cameroon (Diggle et al., 2007). At the regional scale, MBG applications are documented for *S. mansoni* and STHs in East Africa (Clements et al., 2010b), *S. haematobium* in West Africa (Clements et al., 2008c), and lymphatic filariasis in West Africa (Kelly-Hope et al., 2006). Outside Africa, the climatic limits of Asian schistosomiasis caused by *S. japonicum* have been investigated. For example, Wang et al. (2008) used a spatio-temporal model for risk mapping of *S. japonicum* prevalence in the Yangtze River system. Around Lake Dongting in China, Raso et al. (2009a) used MBG to show that the presence of infected buffalos constituted a reservoir of *S. japonicum* and were driving human transmission of this parasite.

Most of these approaches involve logistic regression where the outcome is modelled either as a Bernoulli (Beck-Worner et al., 2007; Raso et al., 2005, 2006a,b, 2009a) or binomial-distributed variable (Clements et al., 2006a, 2008c, 2009, 2010a; Diggle et al., 2007; Kelly-Hope et al., 2006; Wang et al., 2008) depending on whether the data are at the individual level or grouped by location.

Generally, findings of the studies reviewed above have confirmed a geographically focal (i.e. clustered) distribution of helminth infection. However, different helminth infections have different spatial patterns: for example, *S. mansoni* infections in East Africa are clustered near large perennial inland water bodies and hookworm in the same region is relatively widespread within climatically suitable areas.

Figures 5.3-5.6 provide an example of an MBG application using binomial logistic regression models for *S. haematobium*, *S. mansoni* and hookworm prevalence in West Africa (Box 5.4). The resulting maps show that the prevalence of *S. haematobium* infection is much more widely distributed than *S. mansoni* and hookworm in the region and is closely associated with the distance to perennial inland water bodies in Mali and in Ghana (Lake Volta). The covariate coefficients presented in Table 5.1 are consistent with the known epidemiology of schistosome and hookworm infection. The table also shows that clusters of hookworm are smaller than the two schistosome species and there is less propensity for clustering of *S. mansoni* infection compared to *S. haematobium* and hookworm. This contrasts with findings reported in East Africa which show that *S. mansoni* typically has a focal distribution, whereas hookworms are more widely distributed in space (Clements et al., 2006a,b, 2008a,b), and requires further investigation.

5.6.2. Mapping intensity of infection

Spatial modelling of infection intensity can provide additional insight for the design of control programmes not only by identifying high-transmission areas, but by providing a basis for predicting the impact of interventions on morbidity: predictions of infection

intensity can inform the frequency and required coverage of treatment on the basis of mathematical models (Chan et al., 1994, 1995, 1996, 1998).

For schistosomiasis and STH infection, intensity of infection refers to the number of worms in individual hosts and is indirectly measured by quantitative egg counts. For filariasis, intensity of infection is measured by the density of microfilariae from thick blood smears and for onchocerciasis by the density of *O. volvulus* microfilariae in the skin, as assessed by skin snips. Intensity of infection can be represented in a number of ways: mean intensity of infection (regardless of infection status), geometric mean intensity and the prevalence of different categories of infection intensity. To date, the spatial prediction of intensity has been based on multinomial, negative binomial and zero-inflated models, the latter two designed to model overdispersion in individual egg counts. The multinomial approach is the most straightforward and involves predicting the prevalence of low and moderate/heavy intensity infections which can be useful tools for estimating the burden of helminth diseases (Clements et al., 2010a). In Mali, Niger and Burkina Faso, Clements et al. (2010a) used a multinomial formulation to identify areas with the highest prevalence of high-intensity of *S. haematobium* infection and estimated the number of school-age children with high and low intensity infections.

The main limitation of the multinomial approach is that it involves stratifying egg counts, leading to a loss of information, whereas the negative binomial approach makes full use of intensity data on a continuous scale. Therefore, an alternative approach is to model individual level egg counts. In the case of *S. mansoni* or STH infection, this is estimated by the number of eggs per gram of faeces, or the number of eggs per 10 ml urine for *S. haematobium* or density of microfilariae for filariasis (Alexander et al., 2000). Brooker et al. (2006a) showed, using a negative binomial model, that household clustering of heavy intensity infections was more pronounced in rural areas for *S. mansoni* and *A. lumbricoides* but was similar between rural and urban areas for hookworm. The first MBG application for predicting intensity involved fitting a negative binomial distribution to *S. mansoni* intensity data from East Africa to identify environmental factors associated with the spatial heterogeneity in infection intensity and to produce a predictive map (Clements et al., 2006b). This study helped to identify areas where population based morbidity control, using praziquantel, is most warranted and the resulting posterior predictive estimates of infection intensity can be used to model the potential impact of treatment (e.g. by defining the frequency of treatment required to reduce morbidity).

A feature of intensity of infection is that only a small proportion of the infected population excretes large numbers of parasite eggs and therefore intensity data typically contain a majority of zero counts. Therefore, standard Poisson or negative binomial regression models might not be suited for modelling purposes. To address this problem, zero-inflated formulations of the Poisson (ZIP) or negative binomial (ZINB) regression model have been proposed (Filipe et al., 2005; Pion et al., 2006; Vounatsou et al., 2009). Vounatsou et al. (2009) reported the first application of a ZINB model within an MBG framework for *S. mansoni* infection in Cote d'Ivoire. This study showed that the geostatistical zero-inflated models produce more accurate maps of helminth infection intensity than the spatial negative binomial counterparts. Examples of non-spatial applications of such models are available for lymphatic filariasis (Filipe et al., 2005) and loais (Pion et al., 2006) in humans and *Nematodirus battus* in lambs (Denwood et al., 2008).

5.7. METHODOLOGICAL REFINEMENTS IN MODEL-BASED GEOSTATISTICS

Whilst the past decade has seen a dramatic expansion in the number of helminthological studies employing MBG, each incorporating iterative improvements in modelling approach, there remain a number of areas requiring further investigation. Below we highlight three main areas that deserve attention.

5.7.1. Non-stationarity

Most geostatistical predictive maps reported in the literature are based on statistical models that assume stationarity of the spatial process. This means that the covariance of the residuals between any two locations is modelled as dependent on distance and direction between them and is independent of the location itself. While this may be particularly appropriate for small study areas (where spatial processes can be assumed to be approximately stationary), this assumption may not be optimal when considering spatial processes over large geographical areas. A nonstationary model may be more appropriate because of man-made environmental transformations, geographical variation of climate or topography, the implementation of control or different species or strains of parasites, intermediate hosts and vectors, which may drive differing spatial structure from place to place.

The significance of non-stationarity can be assessed by partitioning the study area and observing differences in empirical semi-variograms between the areas. In order to take non-stationarity into account there are several methods that can be implemented in an MBG framework. An early example involved the modelling of a single Gaussian spatial process which varied at increments across regions in a stationary fashion (Kim et al., 2005). An extension of this method to non-Gaussian prevalence data was presented by Gempferli (2003) and involved a Voronoi random tessellation method, using reversible jump MCMC computations, whereby the data ‘choose’ the number and locations of the partitions (or tiles) to be imposed on the region. More recently, researchers have partitioned the study area into disjoint regions, based on arbitrary divisions or ecological zones, and assuming a separate stationary process in each region (Beck-Worner et al., 2007; Raso et al., 2006a,b; Vounatsou et al., 2009). Transition of the autocorrelation functions across regions is smoothed using normalised distance weighted sums. An advantage of this approach is that it requires the inversion of several covariance matrices of smaller dimensions, thus considerably aiding computation when there are a large number of locations. A disadvantage is that the number and division of regions is subjective and the assumption of independence of data across regions questionable.

5.7.2. Incorporating diagnostic uncertainty

The diagnostic sensitivity of a single Kato–Katz thick smear or urine slide examination is low due to significant day-to-day and intra-specimen variation (Utzinger et al., 2001), and low infection intensities are likely to be missed unless multiple samples over consecutive days are collected (Booth et al., 2003; Engels et al., 1996). For STHs, it has been shown that the Kato–Katz technique can perform with reasonable accuracy with one day’s stool collection for *A. lumbricoides* and *T. Trichiura* but not for hookworm (Tarafder et al., 2010). The inclusion of diagnostic uncertainty into modelling is particularly important for schistosomiasis in low transmission settings (Leonardo et al., 2008). Although test sensitivity and specificity are imperfectly measured, plausible values can be incorporated by modelling them as random variables with ‘informative’ priors. Most spatial prediction models for helminth diseases reported thus far have not included diagnostic uncertainty but a spatial prediction model has been recently reported for prevalence of *S. japonicum* in China

(Wang et al., 2008), adjusting for measurement error by modelling true prevalence as a function of the observed prevalence and test sensitivity and specificity, with the generalised linear model fit to the true prevalence parameter.

5.7.3. Non-linear environmental effects

Often the form of the relationship between infection outcome and environmental covariates is non-linear. Non-linearity can be handled parametrically, such as by modelling covariates with polynomials and non-parametrically, such as by using penalised spline regression. A Bayesian approach to penalised spline regression has recently been proposed (Crainiceanu et al., 2005) and demonstrated within an MBG framework (Gosoni et al., 2009). To illustrate the differences in possible approaches, Figs. 5.3 and 5.4 present parametric and penalised spline regression approaches to the modelling of schistosomiasis in West Africa. This approach yielded a risk map which is consistent with the map using the non-spline approach and is smoother than the non-spline counterpart—however, the fit of the spline model to the data is poorer than that of the non-splined model resulting in a higher DIC.

5.8. APPLICATIONS TO PLANNING AND EVALUATING HELMINTH CONTROL

The flexibility afforded by MBG provides a powerful planning tool for the design and implementation of intervention strategies. For schistosomiasis, applications have primarily focused on predicting the prevalence of infection, enabling areas to be stratified according to intervention strategy: for example, identifying areas where the posterior mean predicted prevalence exceeds 50% in Tanzania (Clements et al., 2006a). Possibly more useful for the control programme manager is an estimate of the probability that prevalence exceeds this threshold, using probability contour maps (PCM). In Burkina Faso, Niger and Mali, for example, Clements et al. (2008c) employed MBG to model the probability of prevalence exceeding 50%, the WHO recommended thresholds for MDA. More work needs to be done to communicate the benefits of this probability-based approach to real-world decision-making.

Individuals heavily infected with the helminth *Loa loa* and treated with ivermectin as part of the APOC onchocerciasis control programme are at high risk of potentially fatal encephalopathic adverse reactions. To help identify areas where prevalence of *L. loa* exceeds 20% and increased risk of adverse reactions, Diggle et al. (2007) used MBG to construct a PCM for *L. Loa*, demonstrating areas where infection prevalence exceeds 20% and that require precautionary strategies for managing potential adverse events. Traditionally, uncertainty would be expressed through a map of the prediction variance (or mean square error) but a high prediction variance may or may not translate into a high degree of uncertainty as to whether the intervention threshold is exceeded in a given location. The PCM maps are therefore superior to quantify the strength of the available evidence pointing as to whether the threshold is exceeded.

Helminth control is rarely targeted towards one species alone, and recently there has been increased advocacy for an integrated approach to control, whereby multiple drugs targeting a range of helminth infections are co-implemented in a single programme (Hotez, 2009). To guide integrated control requires information on the geographic overlap of different species (Brooker and Utzinger, 2007; Hotez et al., 2007). A first MBG application of mapping such co-endemicity was provided by Clements et al. (2010a,b) who mapped the co-distribution of *S. mansoni* and one or more soil-transmitted helminths in eastern Africa. Here, hookworm was found to be ubiquitous whilst *S. mansoni* was highly focal, occurring predominantly in locations near the Nile River and the Great Lakes. Therefore, albendazole is required

throughout the region but praziquantel is only required in specific high-risk areas for *S. mansoni*. Figure 5.7 presents the use of a co-endemicity map for the West African Region. This map highlights that areas for twice-annually, integrated MDA for urinary schistosomiasis and hookworm are highly focal across the West African region. A novel approach is mapping co-intensity of parasite infection. Similar to co-endemicity maps, this simply involves overlaying intensity maps for multiple parasite infections on a single map, allowing identification of geographical overlap of areas where transmission of multiple parasites is at its highest. We propose that this mapping approach could be advantageous as a planning and evaluation tool by assisting in geographical targeting of morbidity control and providing an assessment of the progress of successive MDA in integrated programmes.

Where different species overlap in distribution, it is likely that many individuals will harbour co-infections with one or more species. To date, two studies have employed MBG to predict the geographical distribution of parasite co-infections (Brooker and Clements, 2009; Raso et al., 2006a,b). Both studies investigated the spatial distribution of co-infection with *S. mansoni* and hookworm, the first at sub-national scale in Cote D'Ivoire (Raso et al., 2006a,b) and the second at the regional scale in the East African Great Lakes Region (Brooker and Clements, 2009). These studies found that adolescents and males are at increased risk of *S. mansoni* and hookworm co-infections; they also found that the spatial heterogeneities in *S. mansoni* and hookworm co-infections were significantly associated with several environmental covariates (temperature, elevation and distance to large water bodies). In a non-spatial, Bayesian hierarchical modelling study in Brazil, Pullan et al. (2008) found that there was strong evidence of household clustering of *S. mansoni* and hookworm co-infection. These authors found that approximately one-third of the between-household variability was due to socio-economic status, household crowding and high Normalised Difference Vegetation Index (NDVI). All of these studies use multinomial specifications of the outcome and compare mono- and co-infection patterns with no infection.

In addition to mapping the geographical distribution of infection, it is essential for control programmes to establish the total number infected or co-infected, and the population at risk, to estimate resource requirements.

BOX 5.4

General formulation of Bayesian geostatistical models used for producing smooth climate-based maps of helminth diseases

The Bayesian geostatistical models for prevalence were of the form:

$$Y_{ij} \sim \text{Binomial}(n_{ij}, p_{ij})$$

$$\text{logit}(p_{ij}) = \alpha + \sum_{k=1}^p \beta_k \times x_{i,j,k} + u_i$$

where Y_{ij} is the number of infection positive children in school i , age–sex group j , n_{ij} is the number of children examined in school i , age–sex group j , p_{ij} is prevalence of infection in school i , age–sex group j , α is the intercept, x is a matrix of covariates, β is a matrix of coefficients and u_i is a geostatistical random effect defined by an isotropic powered exponential spatial correlation function:

$$f(d_{ab}; \phi) = \exp[-(\phi d_{ab})]$$

where d_{ab} are the distances between pairs of points a and b , and ϕ is the rate of decline of spatial correlation per unit of distance. Non-informative priors were used for α (uniform prior with bounds $-\infty$ and ∞) and the coefficients (normal prior with mean = 0 and precision = 1×10^{-4}). The prior distribution of ϕ was also uniform with upper and lower bounds set at 0.06 and 50. The precision of u_i was given a non-informative gamma distribution.

(Brooker et al., 2006b; Clements et al., 2010a; Tatem et al., 2008). Several electronic population density maps for SSA are freely available on the internet which include the Global Rural-Urban Mapping Project (GRUMP; <http://sedac.ciesin.columbia.edu/gpw/>), the Gridded Population of the World version 3 (GPW3; <http://sedac.ciesin.org/gpw/>), the Landsat 2005 (<http://www.ornl.gov/sci/landscan/>) and, for Kenya, the African Population database (APD; <http://www.na.unep.org/globalpop/africa/>). The GRUMP is a global population distribution map which has a spatial resolution grid of 1 km². It has been demonstrated to be the most accurate of recently available population surfaces (Hay et al., 2005). In GRUMP, sub-national 2000 census data are combined with an urban extent mask that adjusts population totals and densities within areas defined as urban (Balk et al., 2006). Because population gridded products use census datasets, population figures need to be projected to the year of interest. This can be done by using country-specific reported population growth rates available at the United Nations Population Division–World Population Prospectus database (<http://esa.un.org/unpp/>; Brooker et al., 2000a, 2002, 2003, 2006a; Clements et al., 2010a). Predicted prevalence maps, including those derived from MBG, can be multiplied by electronic population density maps to determine the numbers of individuals infected in each location—if the MBG prediction is jointly simulated, these numbers can then be aggregated by administrative area or nationally to determine the overall burden of infection. Additionally, masks can be overlaid on the population density map to delineate areas where transmission does and does not occur and numbers of people at risk can then be calculated.

5.9. CONCLUSION

MBG represents a key advance in the spatial prediction of helminth disease at different spatial scales. There are an increasing number of examples in the published literature where maps produced using these methods have been used in the planning and implementation of disease control programmes. Methods for representing uncertainty constitute a major advantage of MBG compared to classical geostatistics and other spatial prediction methods. However, there is a need to translate the benefits of flexible uncertainty representation in a form readily interpretable to control personnel if MBG is to be maximally utilised. A framework that reconciles control programme objectives, available disease survey data and the various applications within the MBG platform could provide potentially important benefits to current disease control programmes.

Acknowledgments

A. C. A. C. is funded by an Australian National Health and Medical Research Council Career Development Award (#631619), A. P. P. is funded under a Wellcome Trust Principal Research Fellowship held by Professor Bob Snow (#079080), P. W. G. is funded under the Wellcome Trust Senior Research Fellowship held by Dr. Simon Hay (#079091) and S. B. is supported by a Research Career Development Fellowship from the Wellcome Trust (#081673). Finally, we are most grateful to the SCI-supported national programmes in west Africa for allowing us to showcase their survey data in this chapter.

REFERENCES

- Agumya A, Hunter GJ. Responding to the consequences of uncertainty in geographical data. *Int. J. Geogr. Inf. Sci.* 2002; 16:405–417.
- Alexander N, Moyeed R, Stander J. Spatial modelling of individual-level parasite counts using the negative binomial distribution. *Biostatistics*. 2000; 1:453–463. [PubMed: 12933567]
- Balk DL, Deichmann U, Yetman G, Pozzi F, Hay SI, Nelson A. Determining global population distribution: methods, applications and data. *Adv. Parasitol.* 2006; 62:119–156. [PubMed: 16647969]
- Basanez MG, Pion SD, Churcher TS, Breitling LP, Little MP, Boussinesq M. River blindness: a success story under threat? *PLoS Med.* 2006; 3(9):e371. [PubMed: 17002504]
- Beck-Worner C, Raso G, Vounatsou P, N’Goran EK, Rigo G, Parlow E, Utzinger J. Bayesian spatial risk prediction of *Schistosoma mansoni* infection in western Cote d’Ivoire using a remotely-sensed digital elevation model. *Am. J. Trop. Med. Hyg.* 2007; 76:956–963. [PubMed: 17488922]
- Boatin BA, Richards FO Jr. Control of onchocerciasis. *Adv. Parasitol.* 2006; 61:349–394. [PubMed: 16735169]
- Boatin BA, Hougard JM, Alley ES, Akpoboua LK, Yameogo L, Dembele N, Seketeli A, Dadzie KY. The impact of Mectizan on the transmission of onchocerciasis. *Ann. Trop. Med. Parasitol.* 1998; 92:S46–S60. [PubMed: 9861267]
- Booth M, Vounatsou P, N’Goran EK, Tanner M, Utzinger J. The influence of sampling effort and the performance of the Kato-Katz technique in diagnosing *Schistosoma mansoni* and hookworm co-infections in rural Cote d’Ivoire. *Parasitology*. 2003; 127:525–531. [PubMed: 14700188]
- Brooker S. Schistosomes, snails and satellites. *Acta Trop.* 2002; 82:207–214. [PubMed: 12020894]
- Brooker S. Spatial epidemiology of human schistosomiasis in Africa: risk models, transmission dynamics and control. *Trans. R. Soc. Trop. Med. Hyg.* 2007; 101:1–8. [PubMed: 17055547]
- Brooker S, Clements AC. Spatial heterogeneity of parasite co-infection: determinants and geostatistical prediction at regional scales. *Int. J. Parasitol.* 2009; 39:591–597. [PubMed: 19073189]
- Brooker S, Michael E. The potential of geographical information systems and remote sensing in the epidemiology and control of human helminth infections. *Adv. Parasitol.* 2000; 47:245–288. [PubMed: 10997209]
- Brooker S, Utzinger J. Integrated disease mapping in a polyparasitic world. *Geospat. Health.* 2007; 1:141–146. [PubMed: 18686239]
- Brooker S, Donnelly CA, Guyatt HL. Estimating the number of helminthic infections in the Republic of Cameroon from data on infection prevalence in school-children. *Bull. World Health Organ.* 2000a; 78:1456–1465. [PubMed: 11196498]
- Brooker S, Rowlands M, Haller L, Savioli L, Bundy DA. Towards an atlas of human helminth infection in sub-Saharan Africa: the use of geographical information systems (GIS). *Parasitol. Today.* 2000b; 16:303–307. [PubMed: 10858650]
- Brooker S, Hay SI, Tchuente L.-A. Tchuem, Ratard RC. Using NOAA-AVHRR data to model human helminth distributions in planning disease control in Cameroon, West Africa. *Photogramm. Eng. Remote Sens.* 2002; 68:175–179.
- Brooker S, Singhasivanon P, Waikagul J, Supavej S, Kojima S, Takeuchi T, Luong TV, Looareesuwan S. Mapping soil-transmitted helminths in Southeast Asia and implications for parasite control. *Southeast Asian J. Trop. Med. Public Health.* 2003; 34:24–36. [PubMed: 12971511]
- Brooker S, Kabatereine NB, Tukahebwa EM, Kazibwe F. Spatial analysis of the distribution of intestinal nematode infections in Uganda. *Epidemiol. Infect.* 2004; 132(6):1065–1071. [PubMed: 15635963]
- Brooker S, Kabatereine NB, Myatt M, Stothard J, Russell, Fenwick A. Rapid assessment of *Schistosoma mansoni*: the validity, applicability and cost-effectiveness of the Lot Quality Assurance Sampling method in Uganda. *Trop. Med. Int. Health.* 2005; 10:647–658. [PubMed: 15960703]
- Brooker S, Alexander N, Geiger S, Moyeed RA, Stander J, Fleming F, Hotez PJ, Correa-Oliveira R, Bethony J. Contrasting patterns in the small-scale heterogeneity of human helminth infections in

- urban and rural environments in Brazil. *Int. J. Parasitol.* 2006a; 36:1143–1151. [PubMed: 16814294]
- Brooker S, Clements AC, Bundy DA. Global epidemiology, ecology and control of soil-transmitted helminth infections. *Adv. Parasitol.* 2006b; 62:221–261. [PubMed: 16647972]
- Brooker S, Clements AC, Hotez PJ, Hay SI, Tatem AJ, Bundy DA, Snow RW. The co-distribution of *Plasmodium falciparum* and hookworm among African schoolchildren. *Malar. J.* 2006c; 5:99. [PubMed: 17083720]
- Brooker S, Kabatereine NB, Gyapong JO, Stothard JR, Utzinger J. Rapid mapping of schistosomiasis and other neglected tropical diseases in the context of integrated control programmes in Africa. *Parasitology.* 2009; 136:1707–1718. [PubMed: 19450373]
- Brus DJ, Gruijter JJ. Design-based versus model-based estimates of spatial means: theory and applications in environmental soil science. *Environmetrics.* 1993; 4:123–152.
- Brus DJ, Gruijter JJ. Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). *Geoderma.* 1997; 80:1–44.
- Chan MS, Guyatt HL, Bundy DA, Medley GF. The development and validation of an age-structured model for the evaluation of disease control strategies for intestinal helminths. *Parasitology.* 1994; 109:389–396. [PubMed: 7970893]
- Chan MS, Guyatt HL, Bundy DA, Booth M, Fulford AJ, Medley GF. The development of an age structured model for schistosomiasis transmission dynamics and control and its validation for *Schistosoma mansoni*. *Epidemiol. Infect.* 1995; 115:325–344. [PubMed: 7589272]
- Chan MS, Guyatt HL, Bundy DA, Medley GF. Dynamic models of schistosomiasis morbidity. *Am. J. Trop. Med. Hyg.* 1996; 55:52–62. [PubMed: 8702023]
- Chan MS, Srividya A, Norman RA, Pani SP, Ramaiah KD, Vanamail P, Michael E, Das PK, Bundy DA. Epifil: a dynamic model of infection and disease in lymphatic filariasis. *Am. J. Trop. Med. Hyg.* 1998; 59:606–614. [PubMed: 9790439]
- Clements AC, Lwambo NJ, Blair L, Nyandindi U, Kaatano G, Kinung'hi S, Webster JP, Fenwick A, Brooker S. Bayesian spatial analysis and disease mapping: tools to enhance planning and implementation of a schistosomiasis control programme in Tanzania. *Trop. Med. Int. Health.* 2006a; 11:490–503. [PubMed: 16553932]
- Clements AC, Moyeed R, Brooker S. Bayesian geostatistical prediction of the intensity of infection with *Schistosoma mansoni* in East Africa. *Parasitology.* 2006b; 133:711–719. [PubMed: 16953953]
- Clements AC, Barnett AG, Nyandindi U, Lwambo NJ, Kihamia CM, Blair L. Age and gender effects in self-reported urinary schistosomiasis in Tanzania. *Trop. Med. Int. Health.* 2008a; 13:713–721. [PubMed: 18419586]
- Clements AC, Brooker S, Nyandindi U, Fenwick A, Blair L. Bayesian spatial analysis of a national urinary schistosomiasis questionnaire to assist geographic targeting of schistosomiasis control in Tanzania, East Africa. *Int. J. Parasitol.* 2008b; 38:401–415. [PubMed: 17920605]
- Clements AC, Garba A, Sacko M, Toure S, Dembele R, Landoure A, Bosque-Oliva E, Gabrielli AF, Fenwick A. Mapping the probability of schistosomiasis and associated uncertainty, West Africa. *Emerg. Infect. Dis.* 2008c; 14:1629–1632. [PubMed: 18826832]
- Clements AC, Bosque-Oliva E, Sacko M, Landoure A, Dembele R, Traore M, Coulibaly G, Gabrielli AF, Fenwick A, Brooker S. A comparative study of the spatial distribution of schistosomiasis in mali in 1984–1989 and 2004–2006. *PLoS Negl. Trop. Dis.* 2009; 3:e431. [PubMed: 19415108]
- Clements AC, Firth S, Dembele R, Garba A, Toure A, Sacko M, Landoure A, Bosque-Oliva E, Barnett AG, Brooker S, Fenwick A. Use of Bayesian geostatistical prediction to estimate local variations in *Schistosoma haematobium* infection in West Africa. *Bull. World Health Organ.* 2010a; 87:921–929. [PubMed: 20454483]
- Clements ACA, Deville MA, Ndayishimiye O, Brooker S, Fenwick A. Spatial co-distribution of neglected tropical diseases in the East African Great Lakes region: revisiting the justification for integrated control. *Trop. Med. Int. Health.* 2010b; 15:198–207. [PubMed: 20409287]
- Craig MH, Sharp BL, Mabaso ML, Kleinschmidt I. Developing a spatial-statistical model and map of historical malaria prevalence in Botswana using a staged variable selection procedure. *Int. J. Health Geogr.* 2007; 6:44. [PubMed: 17892584]

- Crainiceanu CM, Ruppert R, Wand MP. Bayesian analysis for penalised spline regression using WinBUGS. *J. Stat. Softw.* 2005; 14:1–24.
- Cressie N. The origins of kriging. *Math. Geol.* 1990; 2:239–252.
- Cressie, N. Statistics for spatial data. Wiley; New York: 1993.
- de Silva NR, Brooker S, Hotez PJ, Montresor A, Engels D, Savioli L. Soil-transmitted helminth infections: updating the global picture. *Trends Parasitol.* 2003; 19:547–551. [PubMed: 14642761]
- Denwood MJ, Stear MJ, Matthews L, Reid SW, Toft N, Innocent GT. The distribution of the pathogenic nematode *Nematodirus battus* in lambs is zero-inflated. *Parasitology.* 2008; 135:1225–1235. [PubMed: 18620622]
- Diggle P, Lophaven S. Bayesian geostatistical design. *Scand. J. Stat.* 2006; 33:53–64.
- Diggle P, Tawn J, Moyeed RA. Model-based geostatistics. *Appl. Stat.* 1998; 47:299–350.
- Diggle P, Moyeed R, Rowlingson B, Thomson MC. Childhood malaria in the Gambia: a case-study in model-based geostatistics. *Appl. Stat.* 2002; 51:493–506.
- Diggle PJ, Thomson MC, Christensen OF, Rowlingson B, Obsomer V, Gardon J, Wanji S, Takougang I, Enyong P, Kamgno J, Remme JH, Boussinesq M, Molyneux DH. Spatial modelling and the prediction of *Loa loa* risk: decision making under uncertainty. *Ann. Trop. Med. Parasitol.* 2007; 101:499–509. [PubMed: 17716433]
- Dormann CF. Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Glob. Ecol. Biogeogr.* 2007; 16:129–138.
- Engels D, Nahimana S, Gryseels B. Comparison of the direct faecal smear and two thick smear techniques for the diagnosis of intestinal parasitic infections. *Trans. R. Soc. Trop. Med. Hyg.* 1996; 90:523–525. [PubMed: 8944262]
- Fenwick A, Webster JP, Bosque-Oliva E, Blair L, Fleming FM, Zhang Y, Garba A, Stothard JR, Gabrielli AF, Clements AC, Kabatereine NB, Toure S, Dembele R, Nyandindi U, Mwansa J, Koukounari A. The Schistosomiasis Control Initiative (SCI): rationale, development and implementation from 2002–2008. *Parasitology.* 2009; 136:1719–1730. [PubMed: 19631008]
- Filipe JA, Boussinesq M, Renz A, Collins RC, Vivas-Martinez S, Grillet ME, Little MP, Basanez MG. Human infection patterns and heterogeneous exposure in river blindness. *Proc. Natl. Acad. Sci. USA.* 2005; 102:15265–15270. [PubMed: 16217028]
- Gemperli, A. Doctoral Dissertation. Swiss Tropical Institute, University of Basel; 2003. Development of Spatial Statistical Methods for Modeling Point-Referenced Spatial Data in Malaria Epidemiology; p. 111–127.
- Gemperli A, Vounatsou P, Kleinschmidt I, Bagayoko M, Lengeler C, Smith T. Spatial patterns of infant mortality in Mali: the effect of malaria endemicity. *Am. J. Epidemiol.* 2004; 159:64–72. [PubMed: 14693661]
- Gemperli A, Sogoba N, Fondjo E, Mabaso M, Bagayoko M, Briet OJ, Anderegg D, Liebe J, Smith T, Vounatsou P. Mapping malaria transmission in West and Central Africa. *Trop. Med. Int. Health.* 2006a; 11:1032–1046. [PubMed: 16827704]
- Gemperli A, Vounatsou P, Sogoba N, Smith T. Malaria mapping using transmission models: application to survey data from Mali. *Am. J. Epidemiol.* 2006b; 163:289–297. [PubMed: 16357113]
- Gething PW, Patil AP, Hay SI. Quantifying aggregated uncertainty in *Plasmodium falciparum* malaria prevalence and populations at risk via efficient space-time geostatistical joint simulation. *PLoS Comput. Biol.* 2010; 6:e1000724. [PubMed: 20369009]
- Gosoni L, Vounatsou P, Sogoba N, Smith T. Bayesian modelling of geostatistical malaria risk data. *Geospat. Health.* 2006; 1:127–139. [PubMed: 18686238]
- Gosoni L, Vounatsou P, Sogoba N, Maire N, Smith T. Mapping malaria risk in West Africa using a Bayesian nonparametric non-stationary model. *Comput. Stat. Data Anal.* 2009; 53:3358–3371.
- Gulani A, Nagpal J, Osmond C, Sachdev HP. Effect of administration of intestinal anthelmintic drugs on haemoglobin: systematic review of randomised controlled trials. *BMJ.* 2007; 334:1095. [PubMed: 17434954]
- Gyapong JO, Remme JH. The use of grid sampling methodology for rapid assessment of the distribution of bancroftian filariasis. *Trans. R. Soc. Trop. Med. Hyg.* 2001; 95:681–686. [PubMed: 11816445]

- Gyapong JO, Kyelem D, Kleinschmidt I, Agbo K, Ahouandogbo F, Gaba J, Owusu-Banahene G, Sanou S, Sodahlon YK, Biswas G, Kale OO, Molyneux DH, ROUNGOU JB, Thomson MC, Remme J. The use of spatial analysis in mapping the distribution of bancroftian filariasis in four West African countries. *Ann. Trop. Med. Parasitol.* 2002; 96:695–705. [PubMed: 12537631]
- Hay SI, Noor AM, Nelson A, Tatem AJ. The accuracy of human population maps for public health application. *Trop. Med. Int. Health.* 2005; 10:1073–1086. [PubMed: 16185243]
- Hay SI, Tatem AJ, Graham AJ, Goetz SJ, Rogers DJ. Global environmental data for mapping infectious disease distribution. *Adv. Parasitol.* 2006; 62:37–77. [PubMed: 16647967]
- Hay SI, Guerra CA, Gething PW, Patil AP, Tatem AJ, Noor AM, Kabaria CW, Manh BH, Elyazar IR, Brooker S, Smith DL, Moyeed RA, Snow RW. A world malaria map: *Plasmodium falciparum* endemicity in 2007. *PLoS Med.* 2009; 6:e1000048. [PubMed: 19323591]
- Hooper PJ, Bradley MH, Biswas G, Ottesen EA. The Global Programme to Eliminate Lymphatic Filariasis: health impact during its first 8 years (2000–2007). *Ann. Trop. Med. Parasitol.* 2009; 103:S17–S21. [PubMed: 19843394]
- Hotez PJ. Mass drug administration and integrated control for the world's high-prevalence neglected tropical diseases. *Clin. Pharmacol. Ther.* 2009; 85:659–664. [PubMed: 19322166]
- Hotez P, Raff S, Fenwick A, Richards F Jr, Molyneux DH. Recent progress in integrated neglected tropical disease control. *Trends Parasitol.* 2007; 23:511–514. [PubMed: 17951109]
- Kazembe LN, Kleinschmidt I, Holtz TH, Sharp BL. Spatial analysis and mapping of malaria risk in Malawi using point-referenced prevalence of infection data. *Int. J. Health Geogr.* 2006; 5:41. [PubMed: 16987415]
- Keiser J, Utzinger J. Efficacy of current drugs against soil-transmitted helminth infections: systematic review and meta-analysis. *JAMA.* 2008; 299:1937–1948. [PubMed: 18430913]
- Kelly-Hope LA, Diggle PJ, Rowlingson BS, Gyapong JO, Kyelem D, Coleman M, Thomson MC, Obsomer V, Lindsay SW, Hemingway J, Molyneux DH. Short communication: negative spatial association between lymphatic filariasis and malaria in West Africa. *Trop. Med. Int. Health.* 2006; 11:129–135. [PubMed: 16451336]
- Kim H-M, Mallick BK, Holmes CC. Analyzing nonstationary spatial data using piecewise Gaussian processes. *J. Am. Stat. Assoc.* 2005; 100:653–668.
- Kuhn I. Incorporating spatial autocorrelation may invert observed patterns. *Divers. Distrib.* 2007; 13:66–69.
- Lawson, A. *Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology*. Chapman & Hall/CRC; Boca Raton, FL: 2009.
- Lengeler C, Utzinger J, Tanner M. Questionnaires for rapid screening of schistosomiasis in sub-Saharan Africa. *Bull. World Health Organ.* 2002; 80:235–242. [PubMed: 11984610]
- Leonardo LR, Rivera P, Saniel O, Villacorte E, Crisostomo B, Hernandez L, Baquilod M, Erce E, Martinez R, Velayudhan R. Prevalence survey of schistosomiasis in Mindanao and the Visayas, The Philippines. *Parasitol. Int.* 2008; 57:246–251. [PubMed: 18508406]
- Michael E, Bundy DA. Global mapping of lymphatic filariasis. *Parasitol. Today.* 1997; 13:472–476. [PubMed: 15275135]
- Molyneux DH. Filaria control and elimination: diagnostic, monitoring and surveillance needs. *Trans. R. Soc. Trop. Med. Hyg.* 2009; 103:338–341. [PubMed: 19181357]
- Noma M, Nwoke BEB, Nutall I, Tambala PA, Enyong P, Namsenmo A, et al. Rapid epidemiological mapping of onchocerciasis (REMO): its application by the African Programme for Onchocerciasis Control (APOC). *Ann. Trop. Med. Parasitol.* 2002; 1(96):29–39.
- Noor AM, Clements AC, Gething PW, Moloney G, Borle M, Shewchuk T, Hay SI, Snow RW. Spatial prediction of *Plasmodium falciparum* prevalence in Somalia. *Malar. J.* 2008; 7:159. [PubMed: 18717998]
- Noor AM, Gething PW, Alegana VA, Patil AP, Hay SI, Muchiri E, Juma E, Snow RW. The risks of malaria infection in Kenya in 2009. *BMC Infect. Dis.* 2009; 9:180. [PubMed: 19930552]
- Olsen A. Efficacy and safety of drug combinations in the treatment of schistosomiasis, soil-transmitted helminthiasis, lymphatic filariasis and onchocerciasis. *Trans. R. Soc. Trop. Med. Hyg.* 2007; 101:747–758. [PubMed: 17481681]

- Pion SD, Filipe JA, Kamgno J, Gardon J, Basanez MG, Boussinesq M. Microfilarial distribution of *Loa loa* in the human host: population dynamics and epidemiological implications. *Parasitology*. 2006; 133:101–109. [PubMed: 16764737]
- Pullan RL, Bethony JM, Geiger SM, Cundill B, Correa-Oliveira R, Quinnell RJ, Brooker S. Human helminth co-infection: analysis of spatial patterns and risk factors in a Brazilian community. *PLoS Negl. Trop. Dis.* 2008; 2:e352. [PubMed: 19104658]
- Raso G, Matthys B, N'Goran EK, Tanner M, Vounatsou P, Utzinger J. Spatial risk prediction and mapping of *Schistosoma mansoni* infections among schoolchildren living in western Cote d'Ivoire. *Parasitology*. 2005; 131:97–108. [PubMed: 16038401]
- Raso G, Vounatsou P, Gosoni L, Tanner M, N'Goran EK, Utzinger J. Risk factors and spatial patterns of hookworm infection among schoolchildren in a rural area of western Cote d'Ivoire. *Int. J. Parasitol.* 2006a; 36:201–210. [PubMed: 16259987]
- Raso G, Vounatsou P, Singer BH, N'Goran EK, Tanner M, Utzinger J. An integrated approach for risk profiling and spatial prediction of *Schistosoma mansoni*-hookworm coinfection. *Proc. Natl. Acad. Sci. USA*. 2006b; 103:6934–6939. [PubMed: 16632601]
- Raso G, Li Y, Zhao Z, Balen J, Williams GM, McManus DP. Spatial distribution of human *Schistosoma japonicum* infections in the Dongting Lake Region, China. *PLoS ONE*. 2009a; 4:e6947. [PubMed: 19759819]
- Raso G, Silue KD, Vounatsou P, Singer BH, Yapi A, Tanner M, Utzinger J, N'Goran EK. Spatial risk profiling of *Plasmodium falciparum* parasitaemia in a high endemicity area in Cote d'Ivoire. *Malar. J.* 2009b; 8:252. [PubMed: 19906295]
- Richter J. The impact of chemotherapy on morbidity due to schistosomiasis. *Acta Trop.* 2003; 86:161–183. [PubMed: 12745135]
- Robinson TP. Spatial statistics and geographical information systems in epidemiology and public health. *Adv. Parasitol.* 2000; 47:81–128. [PubMed: 10997205]
- Silue KD, Raso G, Yapi A, Vounatsou P, Tanner M, N'Goran EK, Utzinger J. Spatially-explicit risk profiling of *Plasmodium falciparum* infections at a small scale: a geostatistical modelling approach. *Malar. J.* 2008; 7:111. [PubMed: 18570685]
- Simoonga C, Utzinger J, Brooker S, Vounatsou P, Appleton CC, Stensgaard AS, Olsen A, Kristensen TK. Remote sensing, geographical information system and spatial analysis for schistosomiasis epidemiology and ecology in Africa. *Parasitology*. 2009; 136:1683–1693. [PubMed: 19627627]
- Srividya A, Michael E, Palaniyandi M, Pani SP, Das PK. A geostatistical analysis of the geographic distribution of lymphatic filariasis prevalence in southern India. *Am. J. Trop. Med. Hyg.* 2002; 67:480–489. [PubMed: 12479548]
- Steinmann P, Keiser J, Bos R, Tanner M, Utzinger J. Schistosomiasis and water resources development: systematic review, meta-analysis, and estimates of people at risk. *Lancet Infect. Dis.* 2006; 6:411–425. [PubMed: 16790382]
- Sturrock HJ, Picon D, Sabasio A, Oguttu D, Robinson E, Lado M, et al. Integrated mapping of neglected tropical diseases: epidemiological findings and control implications for northern Bahr-el-Ghazal State, Southern Sudan. *PLoS Negl. Trop. Dis.* 2009; 3(10):e537. [PubMed: 19859537]
- Takougang I, Meremikwu M, Wandji S, Yenshu EV, Aripko B, Lamle SB, Eka BL, Enyong P, Meli J, Kale O, Remme JH. Rapid assessment method for prevalence and intensity of *Loa loa* infection. *Bull. World Health Organ.* 2002; 80:852–858. [PubMed: 12481206]
- Tarafder MR, Carabin H, Joseph L, Balolong EJ, Olveda R, McGarvey ST. Estimating the sensitivity and specificity of Kato-Katz stool examination technique for detection of hookworms, *Ascaris lumbricoides* and *Trichuris trichiura* infections in humans in the absence of a 'gold standard'. *Int. J. Parasitol.* 2010; 40:399–404. [PubMed: 19772859]
- Tatem AJ, Guerra CA, Kabaria CW, Noor AM, Hay SI. Human population, urban settlement patterns and their impact on *Plasmodium falciparum* malaria endemicity. *Malar. J.* 2008; 7:218. [PubMed: 18954430]
- Taylor-Robinson DC, Jones AP, Garner P. Deworming drugs for treating soil-transmitted intestinal worms in children: effects on growth and school performance. *Cochrane Database Syst. Rev.* 2007; (4):CD000371. <http://www2.cochrane.org/reviews/en/ab000371.html>. [PubMed: 17943740]

- Thomson MC, Obsomer V, Kamgno J, Gardon J, Wanji S, Takougang I, Enyong P, Remme JH, Molyneux DH, Boussinesq M. Mapping the distribution of *Loa loa* in Cameroon in support of the African Programme for Onchocerciasis Control. *Filaria J.* 2004; 3:7. [PubMed: 15298709]
- Tobler WR. A computer movie simulating urban growth in the Detroit region. *Econ. Geogr.* 1970; 46:234–240.
- Utzinger J, Booth M, N’Goran EK, Muller I, Tanner M, Lengeler C. Relative contribution of day-to-day and intra-specimen variation in faecal egg counts of *Schistosoma mansoni* before and after treatment with praziquantel. *Parasitology.* 2001; 122:537–544. [PubMed: 11393827]
- Vounatsou P, Raso G, Tanner M, N’Goran EK, Utzinger J. Bayesian geostatistical modelling for mapping schistosomiasis transmission. *Parasitology.* 2009; 136:1695–1705. [PubMed: 19490724]
- Wackernagel, H. Multivariate geostatistics: an introduction with applications. Springer; Berlin: 2003.
- Waller, LA. Optimal spatial design. In: El-Shaarawi, A.; Piegorsch, W., editors. *Encyclopaedia of Environmetrics*. Vol. Vol. 3. Wiley; New Jersey: 2002. p. 2067-2073.
- Wang XH, Zhou XN, Vounatsou P, Chen Z, Utzinger J, Yang K, Steinmann P, Wu XH. Bayesian spatio-temporal modeling of *Schistosoma japonicum* prevalence data in the absence of a diagnostic ‘gold’ standard. *PLoS Negl. Trop. Dis.* 2008; 2:e250. [PubMed: 18545696]
- WHO. Prevention and Control of Schistosomiasis and Soil-Transmitted Helminthiasis. World Health Organization; Geneva: 2002. p. 57WHO Technical Report Series 912
- WHO. Preventive Chemotherapy in Human Helminthiasis. World Health Organization; Geneva: 2006.

FIGURE 5.1.

Framework for spatial planning and evaluation of parasitic disease control programmes that include (A) integration of single disease control programmes, (B) identification of resource needs and intervention coverage and (C) integration evaluation tools such as mathematical models of disease dynamics and economic evaluation methods. Dashed line, limited use; full line, potential use.

FIGURE 5.2.

Raw prevalence of (A) *Schistosoma haematobium*, (B) hookworm, (C) *Schistosoma mansoni* in school-aged children, West Africa, 1998–2005.

FIGURE 5.3.
 Predicted prevalence of *Schistosoma haematobium* infection in boys aged 15–19, West Africa 1998–2007. Estimates are the mean posterior predicted prevalence values from Bayesian geostatistical models.

FIGURE 5.4.

(A) Predicted prevalence of *Schistosoma haematobium* infection in boys aged 15–19, West Africa (inset Ghana), 1998–2007 based on a spline model. Estimates are the mean posterior predicted prevalence values from Bayesian geostatistical models. (B) Estimated non-linear effect of environmental factors on *S. haematobium* risk in West Africa, based on the P-spline model. The posterior mean probability of infection (full line) and the 95% credible interval are shown.

FIGURE 5.5.

Predicted prevalence of *Schistosoma mansoni* infection in boys aged 15–19, East Africa (inset Ghana), 1998–2007. Estimates are the mean posterior predicted prevalence values from Bayesian geostatistical models.

FIGURE 5.6. Predicted prevalence of hookworm infection in boys aged 15–19, West Africa (inset Ghana), 1998–2007. Estimates are the mean posterior predicted prevalence values from Bayesian geostatistical models.

FIGURE 5.7.

Predicted areas of co-endemicity for *Schistosoma haematobium*, and hookworm in West Africa. For both infections, non-endemic is defined as prevalence <10%, endemic is defined as prevalence >0.1–0.5 and hyper-endemicity is defined as prevalence >50%.

TABLE 5.1

Odds ratios and spatial effects for prevalence of *Schistosoma haematobium* and soil-transmitted helminths (hookworm, *Ascaris lumbricoides* and *Trichuris trichiura*) infections in schoolchildren in Mali, Niger, Burkina Faso and Ghana, 2004–2008

Variable	<i>Schistosoma haematobium</i> Posterior mean (95% CI)	Hookworm Posterior mean (95% CI)	<i>Schistosoma mansoni</i> Posterior mean (95% CI)
Female (vs. male)	0.75 (0.70–0.80)	0.32 (0.23–0.66)	0.58 (0.48–0.70)
Age 15–19 years (vs. 5–9 years)	1.48 (1.35–1.61)	1.78 (1.54–2.13)	1.89 (1.54–2.31)
Distance to PIWB *	0.56 (0.29–1.01)	1.76 (0.67–2.36)	0.81 (0.23–1.97)
Land surface temperature *	0.63 (0.28–1.11)	0.13 (0.09–0.56)	0.37 (0.04–1.20)
Land surface temperature ² *	0.70 (0.50–0.97)	0.50 (0.23–0.78)	0.74 (0.36–1.27)
Intercept	0.10 (0.04–0.25)	0.66 (0.24–0.87)	0.0001 (0.00003– 0.0004)
ρ (rate of decay of spatial correlation)	1.52 (1.01–2.09)	2.38 (2.12–3.78)	2.02 (0.82–4.25)
σ^2 (variance of spatial random effect)	0.12 (0.08–0.15)	0.16 (0.11–0.25)	0.08 (0.01–0.15)

* Variables were standardised to have mean = 0 and standard deviation 1; CI, Bayesian credible interval; PIWB, perennial inland water body; NDVI, normalised difference vegetation = index; 2, Land surface temperature squared.