Spatial and temporal distribution of Onchocerciasis in West Africa with particular reference to Ghana

Simon O'Hanlon

00476924

Thesis submitted for PhD examination

Department of Infectious Disease Epidemiology

Imperial College London

2016

Abstract

Background. Quantifying and predicting the effect of control interventions against infectious diseases requires knowledge about prevalence prior to the implementation of such control measures as well as knowledge of key parameters which affect model outcomes. Here is presented the first geostatistical map estimating the prevalence of human onchocerciasis in the former Onchocerciasis Control Programme in West Africa (OCP) before the initiation of vector control operations in 1975. The map is used to quantify the burden of onchocercal skin disease prior to interventions. The results of a socioeconomic survey investigating coverage and compliance to community-directed treatment with ivermectin are then presented.

Methods and Findings. The OCP epidemiological database was investigated for spatial autocorrelation. A total of 737 village surveys had parasitological examinations conducted prior to the start of control. Using these 737 pre-control data points plus environmental covariates, a Bayesian model-based geostatistical (B-MBG) approach was used to generate a continuous surface (at a pixel resolution of 25km^2) of microfilarial prevalence in West Africa prior to the commencement of the OCP. The mean Pearson's correlation between observed prevalence and model-estimates of prevalence at hold-out locations was 0.693; mean prediction error was 0.77% and mean absolute prediction error was 12.2%. Within OCP boundaries, 17.8 million people were deemed to be at risk and 7.6 million infected in 1975 which is greater than previous estimates. The mean microfilarial prevalence across all countries was 45% (range: 2–90%).

Conclusions and Significance. This is the first map of continuous onchocerciasis prevalence in West Africa, developed using robust geostatistical methods. Important environmental predictors of infection prevalence were identified and used in a model that out-performs those without spatial-random effects or environmental covariates. Our results may be compared with recent epidemiological mapping efforts to find areas of persisting transmission, informing feasibility of elimination with current and novel tools.

Declaration of originality

I declare that the work presented here is my own under the supervision of Professor Maria-Gloria Basáñez and Professor Robert Cheke. All prior work has been appropriately referenced.

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Acknowledgements

First and foremost, I would like to thank my family, Amelia, Sonny and Margot. For putting up with me through producing this thesis, and for the sacrifices you guys made to help me get it finished (especially my long-suffering wife). I love you all. I would also like to thank my parents for their support, especially with baby-sitting.

I owe a deep debt of gratitude to my supervisor Professor Maria-Gloria 'MG' Basáñez. First, for the opportunity to work with her on such an exciting topic, but also for her generosity with her time, her patience, support and her encouragement, particularly when things were not going as smoothly as one might hope. I would also like to pay thanks to my secondary supervisor Professor Robert 'Bob' Cheke. His support, encouragement and help, kept me sane and drag me toward the finishing line.

Finally, I would like to thank all friends and colleagues in DIDE and elsewhere who have helped me in any way. I am grateful to everyone, and in particular to Hannah Slater, with whom I had my first conversations about geostatistical mapping.

Table of Contents

Abstract **2**

Declaration of originality **3**

Acknowledgements **4**

Table of Contents **5**

Table of figures **8**

List of tables **13**

List of Abbreviations **16**

Chapter 1. General Introduction **17**

1.1 Opening Remarks 18

1.2 Lifecycle and detection of the parasite 18

1.3 Diagnosing cases of onchocerciasis 22

1.4 Clinical manifestations of onchocerciasis 24

1.5 The *Simuliidae* vector hosts of onchocerciasis 27

1.6 Onchocerciasis intervention programmes 28

1.7 Intervention Strategies 32

1.8 Mapping of human onchocerciasis 35

1.9 Thesis aims and objectives 40

Chapter 2. Investigating and modelling spatial autocorrelation in Onchocerciasis Control Programme data **44**

2.1 Overall aims and objectives 45

2.2 Introduction 45

2.3 Methods 46

2.4 Results 72

2.5 Conclusions 96

Chapter 3. Bayesian geostatistical modelling of the prevalence of human onchocerciasis using environmental covariates 98

3.1 Chapter Aims And Objectives 99

3.2 Introduction 99

3.3 Methods 100

3.4 Results 115

3.5 Discussion 134

Chapter 4. Estimating the number of persons infected and at-risk of onchocerciasis prior to the onset of the Onchocerciasis Control Programme 142

4.1 Introduction 143

4.2 Methods 147

4.3 Results 155

4.4 Discussion 160

Chapter 5. Socioeconomic analysis of factors associated with onchocerciasis and compliance to community-directed treatment with ivermectin 164

5.1 Chapter Aims and Objectives 165

5.2 Introduction 165

5.3 Methods 167

5.4 Results 174

5.5 Conclusions 209

5.6 Acknowledgements 216

Chapter 6. General Discussion 217

6.1 Summary of work in this thesis 218

6.2 From control to elimination of NTDs 219

6.3 Model-based Bayesian geostatistical mapping 219

6.4 Disease Burden Estimation 226

6.5 Coverage and compliance with community-directed treatment with ivermectin 230

6.6 Future research directions 233

References 235

Appendix A. Vectors of Onchocerciasis 254

Appendix B. Historical Maps of Onchocerciasis Prevalence in West Africa 255

Appendix C. Processing of environmental covariates 263

C.1. Sampling covariate layers at data locations 263

Appendix D. Socioeconomic Survey Instrument 272

Appendix E. Literature review for morbidity and prevalence data 283

Appendix F. Papers Published During PhD Registration 288

F.1. Paper from Chapter 3 288

F.2. Papers not directly related to PhD 289

Table of figures

Figure 1-1: Life-cycle of Onchocerca volvulus. 19

Figure 1-2: The area covered by the Onchocerciasis Control Programme (OCP). 30

Figure 1-3: Global distribution of onchocerciasis. 31

Figure 1-4: Operational phases of the OCP. 33

Figure 2-1: Location of villages in the OCP epidemiological database. 48

Figure 2-2: Location of pre-control villages selected for use in this study. 51

Figure 2-3: Endemicity class of pre-control villages. 52

Figure 2-4: Euclidean distance to nearest river. 54

Figure 2-5: Digitised species identification map. 55

Figure 2-6: Voronoi tessellations of Simulium species. 56

Figure 2-7: Plots of data coordinates against microfilarial prevalence. 72

Figure 2-8: Mantel correlograms of spatial autocorrelation. 73

Figure 2-9: Spatial weight graph constructions. 75

Figure 2-10: Local Moran's I scatterplots. 77

Figure 2-11: Locations of points exhibiting significant local clustering. 78

Figure 2-12: Polynomial trend surface of microfilarial prevalence. 80

Figure 2-13: Multilevel B-spline surfaces of mf prevalence. 81

Figure 2-14: Variogram clouds with theoretical variogram model fits. 87

Figure 2-15: Binned sample semivariograms with fitted theoretical variogram models. 88

Figure 2-16: Ordinate weighting and spatial lag bias. 89

Figure 2-17: Simple kriging surface of mf prevalence. 90

Figure 2-18: Ordinary kriging surfaces with locally estimated mean 91

Figure 2-19: Maps of microfilarial prevalence estimated using external trend kriging. 92

Figure 2-20: Linear correlation between observed and estimated mf prevalence using randomly partitioned validation data 93

Figure 2-21: Linear correlation between observed and estimated mf prevalence using spatially clustered validation data. 94

Figure 3-1: Variograms of spatial trends. 105

Figure 3-2: The location of held-out validation points used in each K-fold cross-validation dataset. 113

Figure 3-3: Highest posterior density plots for fixed effects. 117

Figure 3-4: MCMC traceplot of the value of the range parameter. 118

Figure 3-5: MCMC traceplots of all β_j regression coefficients. 118

Figure 3-6: The pre-control spatial distribution of onchocerciasis microfilarial prevalence in west Africa. 119

Figure 3-7: Alternative depiction of the pre-control distribution of onchocerciasis microfilarial prevalence in west Africa 120

Figure 3-8: Endemicity class of microfilarial prevalence in west Africa in 1975. 121

Figure 3-9: Histogram of mean and median microfilarial prevalence at all prediction locations. 122

Figure 3-10: Maps of the probability of threshold exceedance. 124

Figure 3-11: Probability of hyperendemic threshold exceedance and endemicity class membership. 126

Figure 3-12: Maps of the standard deviation and variance of predictive posteriors..... 128

Figure 3-13: Population weighted index of uncertainty. 129

Figure 3-14: Signal to noise ratio of mean and standard deviation of predictive posteriors. 130

Figure 3-15: Scatter-plot of observed versus estimated microfilarial prevalence (P_{mf}) at validation locations. 131

Figure 3-16: Probability-probability plot..... 132

Figure 3-17: Frequency histogram of observed and estimated P_{mf} at validation locations..... 133

Figure 4-1: Example output from random effect generalised linear model for acute papular onchodermatitis (APOD). 150

Figure 4-2: Estimated log10+1 transformed population counts for 1975..... 152

Figure 4-3: Gridded rural-urban extents mask..... 152

Figure 4-4: Age specific prevalence of APOD. 159

Figure 4-5: Age specific prevalence of DPM. 159

Figure 4-6: Age specific prevalence calculated from the chronic papular onchodermatitis (CPOD) model..... 159

Figure 5-1: Daily activities of a blind man from Asubende. 167

Figure 5-2: Comparison of REMO maps from Ghana Health Service surveys, with end of programme OCP survey data. 169

Figure 5-3: Map of selected village locations. 170

Figure 5-4: Photographs of the interview team at work on this study. 172

Figure 5-5: Example of the CSPro data entry program. 173

Figure 5-6: Age distribution of study population. 176

Figure 5-7: Total number of problems listed in relation to work, sorted by village. 181

Figure 5-8: Age distribution employed persons 186

Figure 5-9: Box-plots showing the distribution of values of different types of livestock (per animal). 187

Figure 5-10: Count of persons successfully completing each school year. 189

Figure 5-11: Distribution of responses about how the main survey respondent felt about different social groups. 191

Figure 5-12: Distribution of number of persons per household who report suffering onchocerciasis, where at least one person reports they suffer from onchocerciasis. 193

Figure 5-13: The number of treatments missed by an individual according to their recall of when they last took ivermectin, the first time they took ivermectin and the number of treatments in total that they received. 197

Figure 5-14: An example of some of the treatment records that were recovered whilst on site. 198

Figure 5-15: Comparison of pure respondent recall of number of treatments they received, vs, the number of treatments actually administered to each village. 199

Figure 5-16: Typical scene of a mudwall home, with thatched roof and external area for cooking. In many cases, livestock are free range and roam freely around villages. 203

Figure 5-17: Summary pages from treatment registers at Jagbenbendo. 211

Figure 5-18: My outline sketches of the villages of Agblekeme and Asubende. 213

Figure 5-19: District level coverage record for Mandari sub-district, Bole district, Northern region.215

Figure 6-1: Recent updates in population distribution maps. 223

Figure 6-2: Screenshot of http://www.ntdmap.org. 225

Figure 6-3: Variation in the relationship between prevalence and intensity of infection with mf. 228

Figure 6-4: Simulation of 30 rounds of CDTI using EPIONCHOv2 (*279*) under different assumptions of systematic non-compliance 231

List of tables

Table 2-1: Distribution of pre-control points across OCP phases. 50

Table 2-2: Global Moran's I spatial autocorrelation statistics..... 74

Table 2-3: Validation statistics from multilevel B-spline surface fitting. 81

Table 2-4: Maximised log-likelihood values for theoretical semivariogram models fitted to empirical semivariograms. 83

Table 2-5: Minimised value of weighted least-squares for theoretical semivariograms models fitted to empirical semivariograms. 84

Table 2-6: Theoretical semivariogram parameter estimates from exponential correlation function.... 86

Table 2-7: Validation statistics for kriging models incorporating an external trend. 95

Table 3-1: Sources of spatially referenced data considered for use as explanatory covariates. 103

Table 3-2: Posterior parameter values. 116

Table 3-3: Observed versus estimated endemicity class classification table..... 132

Table 4-1: Country-specific medium-variant intercensal growth rates. 154

Table 4-2: Total population, at-risk population, numbers with onchocerciasis and mean microfilarial prevalence by OCP country in 1975, at the commencement of the program. 156

Table 4-3: Total number of cases of onchocercal skin disease by country in 1975. 157

Table 4-4: Total number of healthy years of life lost due to years living with disease of onchocercal skin disease (OSD) in the rural popualtion aged 5+ in 1975. 158

Table 5-1: Overall treatment coverage for 7 years prior to this study. 174

Table 5-2: Marital status of all study participants aged 12 years and older. 176

Table 5-3: Household status of household members from 280 households across the two regions in the study. 177

Table 5-4: Number and proportion of people within each study region, who observe a religion. 178

Table 5-5: Table displaying median pay and the number of people employed by economic sector in the study population. 180

Table 5-6: Total number of problems related to work experienced by people across different localities. 181

Table 5-7: Table of average value assigned to tools and equipment, shown by village and region... 183

Table 5-8: Table of distribution of farming tools across villages and regions. 183

Table 5-9: Table of average value assigned to household possessions and other assets, shown by village and region. 185

Table 5-10: Number of possessions owned by village and region, for the most popular categories.. 185

Table 5-11: Number and type of animals kept by household across the study region. 187

Table 5-12: Number and type of qualifications achieved by school leavers, stratified by study region. 190

Table 5-13: Distribution of various household characteristics related to socioeconomic setting and access to services. 192

Table 5-14: Distribution of self-reported prevalence of sysmtpoms related to infection with onchocerciasis. 194

Table 5-15: Types of complier stratified by self-reported onchocerciasis infection status. 195

Table 5-16: Results of a binary logistic regression of compliance status against self-reported infection status. 195

Table 5-17: Proportion of persons stratified by compliance type within each village. 196

Table 5-18: Village knowledge about CDTI interventions..... 200

Table 5-19: Individual level demographic factors and onchocerciasis status..... 201

Table 5-20: Household level demographic factors and onchocerciasis status..... 204

Table 5-21: Individual level sociodemographic variables and their association with recent adherence to CDTI. 206

Table 5-22: Household level sociodemographic variables and their association with recent adherence to CDTI. 207

Table 5-23: Independent predictors of compliance included in a fully adjusted logistic regression model..... 208

List of Abbreviations

APOC, African Programme for Onchocerciasis Control

APOD, Acute Papular Onchodermatitis

ATP, annual transmission potential

BCI, Bayesian credible interval

B-MBG, Bayesian model-based geostatistics

CDTI, community-directed treatment with ivermectin

CPOD, Chronic Papular Onchodermatitis

CV, cross validation

DALY, Disability Adjusted Life Year

DIC, deviance information criterion

DPM, Depigmentation

DRG4, WHO/TDR Disease Reference Group on Helminth Infections

ESPEN, Expanded Special Project for the Elimination of NTDs

EVI, enhanced vegetation index

GLMM, generalised linear mixed model

GLSM, generalised linear spatial model

GRUMP, Gridded Rural Urban Mapping Project

MAP, Malaria Atlas Project

mf, microfilarial

MBG, model-based geostatistics

MCMC, Markov Chain Monte Carlo

NDVI, normalized difference vegetation index

NTD, neglected tropical disease

OCP, Onchocerciasis Control Programme in West Africa

OSD, Onchocercal Skin Disease

PER, Parasite Establishment Rate

REMO, rapid epidemiological mapping of onchocerciasis

***S. damonsum* s.s./s.l.**, *Simuliuum danosum* sensu stricto/sensu lato

SOR, Sub-optimally responding

TDR, UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases

UNDP, United Nations Development Programme

UNICEF, United Nations International Children's Emergency Fund

WHO, World Health Organization.

Chapter 1. General Introduction

1.1 Opening Remarks

Human onchocerciasis (river blindness) is a parasitic infection caused by the filarial nematode *Onchocerca volvulus*. Onchocerciasis belongs to the group of neglected tropical diseases (NTDs). In January 2012, the London Declaration on NTDs committed to sustain, expand and extend the necessary supply of drugs and other interventions to help control onchocerciasis by 2020 (*1*). The life-cycle of *O. volvulus* includes transmission by biting flies of the genus *Simulium*. Persistent infection can lead to gradual loss of visual acuity, eventually leading to blindness, and to the development of a variety of severe and debilitating skin lesions (*2, 3*) through the production of microfilariae (mf), the embryonic stages of the parasite, by adult female worms. Most of the pathology associated with onchocerciasis is due to inflammatory host responses to dying microfilariae in the eyes or skin. The infection is also a cause of excess mortality of the human host, both indirectly through blindness and directly through poorly understood pathogenic mechanisms caused by heavy infection in sighted individuals (*4, 5*).

1.2 Lifecycle and detection of the parasite

1.2.1 Lifespan and localization of adult worms

Onchocerca volvulus is a dioecious (filarial nematode) species (i.e. with separate sexes), necessitating the presence of both male and female worms to produce the microfilariae. It has been estimated that adult females have an average life expectancy of 9-14 years (a mean lifespan of 10 years with a 95% percentile of life expectancy of 13.7 years (*6*). The adult worms are found in nodules, known as onchocercomata, a proportion of which lie within the subcutaneous tissue of the skin and are palpable. A further proportion are also found in more deep-lying nodules in muscular and connective tissues and are more difficult to detect (*7*).

It is assumed that the nodules are formed as a host response to the continuous production of filarial proteins (*8*). As proteins are released, leucocytes and fibroblasts are recruited to the site, thereby forming a granuloma around the female worms that remain sessile within these nodules for the duration of their lives (*7*). By contrast, male worms are assumed to be able to migrate from nodule to nodule, fertilising females and giving rise to a polygamous mating system (*9*). In theory, all female worms within a human host could be fertilised by a single motile male. A single nodule may contain up to 50 adult worms, although on average, 80% of onchocercomata contain one or two male worms and two or three female worms (*7*).

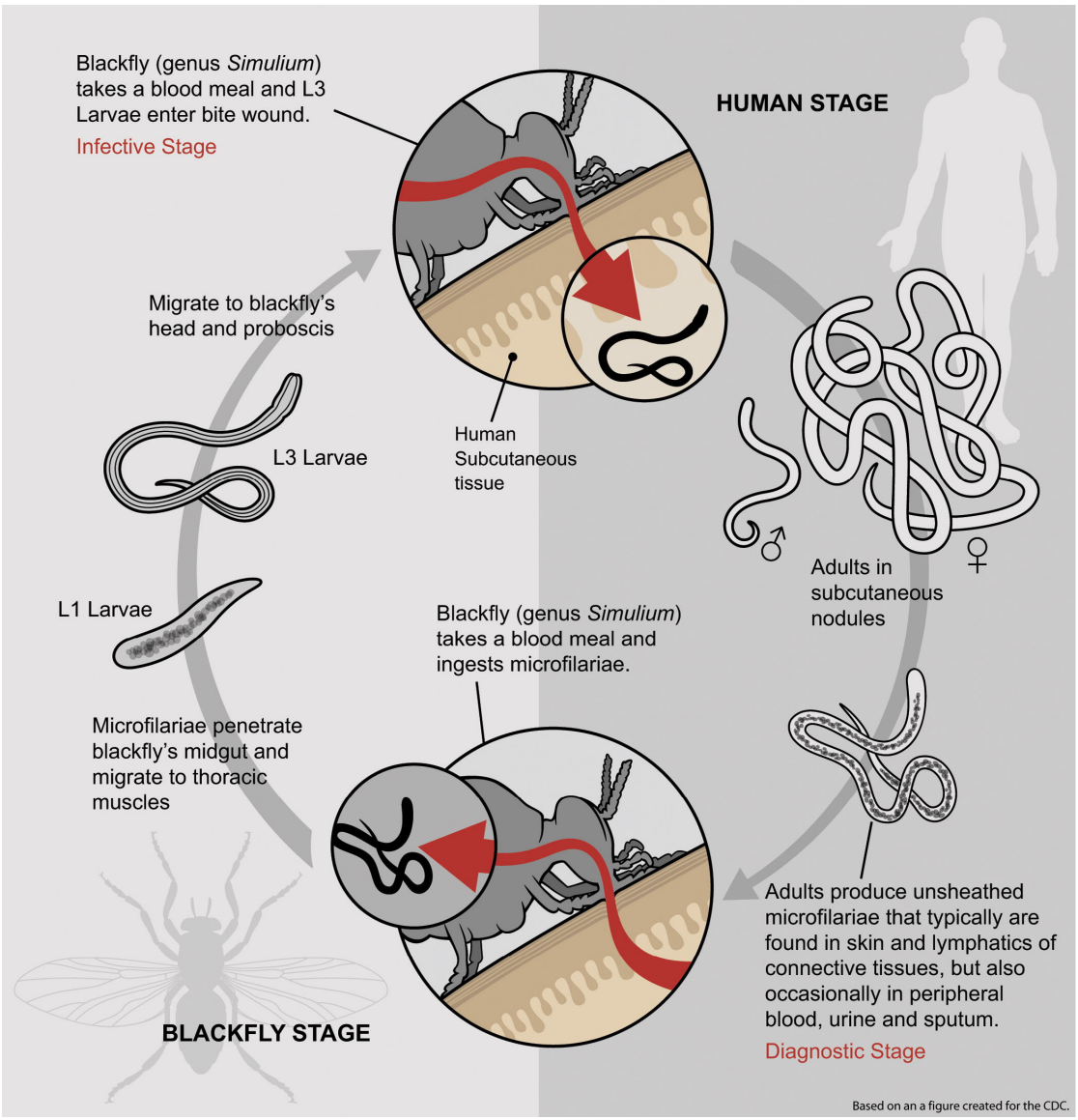


Figure 1-1: Life-cycle of Onchocerca volvulus.
From (*10*). Illustration: Giovanni Maki, derived from a CDC image at http://www.dpd.cdc.gov/dpdx/HTML/Filariasis.htm.

1.2.2 Development from microfilaria to L1 larvae

Fertilised female *O. volvulus* produce between 500 and 1500 *mf* per day (*11*), which have a lifespan of up to 30 months (*12*), and an average between 12 and 15 months (*13*). Heavily infected individuals may experience a cumulative microfilarial load in the order of hundreds of millions of *mf* over the course of their lifetime, and an individual infected with even a single (mated) female *O. volvulus* can carry a load of more than one million microfilariae at any one time.

Importantly, microfilariae of *O. volvulus* must leave the female one by one, i.e. *mf* are not expelled *en*-masse. This process takes 5-10 seconds per *mf* and is therefore a limiting factor (*7*). Once produced *mf* migrate toward, and invade the skin where they continue their lifecycle via an obligate haematophagous simuliid vector after their uptake during a bloodmeal.

1.2.3 Development from microfilaria to L3 infective larvae within vector hosts

Once within their insect hosts, *mf* develop into an L1 stage larva and then via an L2 larval stage to the L3 infective stage, which within *Simulium* species takes between 6 and 12 days (*12*). The ingested *mf* migrate towards the thoracic muscles where larval development takes place (*14*). The *mf* do not multiply within the vector so the L3 infective load within *Simulium* is directly (but not necessarily linearly) determined by the *mf* uptake and any subsequent regulation of parasite numbers within the vector (*15-17*).

1.2.4 Development from L3 larvae to adult worms in human hosts

An uninfected human host may be inoculated with L3 infective larvae when an infective female blackfly takes a blood meal. L3-stage larvae inoculated into human hosts are sexually immature and therefore also unable to reproduce. Development of heavy infections will only occur in those

individuals exposed to many bites from infectious *Simulium* flies over a number of years in the absence of microfilaricidal interventions (see *Section 1.6, Intervention Programmes*). A proportion of developing L3 larvae moult 3-7 days post-inoculation into L4 larvae, and then again 4-6 weeks later into the juvenile adult stage (L5) (*18, 19*). In those hosts with an existing infection, a proportion of immature worms are attracted to existing nodules via an unknown mechanism, forming satellite or composite nodules (*20*). It will take up to a year for a newly acquired worm to mature fully to reproductive capacity, and somewhat longer for the mf to reach detectable levels in the skin (see *Section 1.3 Diagnostics of Onchocerciasis*) with current methods.

1.2.5 Forest and savannah parasite-vector species complexes

Two 'strains' or 'forms 'of *O. volvulus* can cause disease in humans. The most serious ocular pathology (i.e. blindness) is mainly associated with infection by parasites from the savannah regions of Sub-Saharan Africa (Sudan-type and Guinea-type of savannah), transmitted by the so-called savannah members of the *S. damnosum* complex (*S. damnosum* sensu stricto (s.s.) and *S. sirbanum*) (*21-23*). Milder, non-blinding pathology, tends to be associated with endemic infection in forested regions of West Africa, transmitted by the forest species of the complex (e.g. *S. yahense*) (*24*).

Prior to control interventions and in savannah regions, the prevalence of blindness was proportional to the intensity of infection found in those communities affected (*25*). The different pathologies in these regions are thought to be due to differences in the local parasite populations rather than the dynamics of infection itself. DNA-probe based analyses on strains revealed in the two different endemic regions found a strong correlation between disease severity and *O. volvulus* found in savannah areas (*26*). However, the situation is probably more complex than a simple dichotomy between savannah and forest, as blindness rates in forest-savannah mosaic and other types of forest can also be very high (*27*).

1.3 Diagnosing cases of onchocerciasis

1.3.1 Skin snipping

Methods based on skin snipping have traditionally been the gold standard diagnostic for determining infection with *O. volvulus mf*. A small biopsy of skin is taken using a specialised (typically corneoscleral) punch (with a given diameter to provide snips of a relatively uniform size and weight) from parts of the body deemed to have a high concentration of *mf* according to disease ecology, habits of biting vectors (preferentially biting on upper or lower parts of the human body), and studies recording the distribution of microfilariae in human skin (e.g. iliac crests in Africa) (*28*). Persons are counted as positive for onchocerciasis if *mf* emerge from the snips after their incubation in an appropriate solution medium (saline or other) for up to 24h (*29*).

Frequently (but not always), the numbers of *mf* emerging from the skin snip are counted to give an indication of intensity of infection (*mf* per snip) and this was the method employed throughout the Onchocerciasis Control Programme in West Africa (OCP) (see section 1.6: *Onchocerciasis Intervention Programmes*). Skin snipping results (mostly from only two snips taken from the iliac crests in Africa as described above) are commonly presented as either prevalence (the proportion of examined individuals with microfilaridermia) in the population aged five years and above (fraction of individuals with at least one microfilaria in a skin snip) or intensity, the geometric mean load in the population aged 20 years and above (community microfilarial load or CMFL (*30*)).

Skin snipping, though sensitive and specific, is time consuming, invasive and expensive to carry out at a large scale. However, it provides the only measurement of intensity of infection available to date. It is easy to see, however, that considerable measurement error will be present when measuring infection intensity as described above. Furthermore, during the OCP, those skin snips that resulted in a positive result after 30 min incubation (in distilled water) were not incubated further, leading to underestimates of the numbers of microfilariae present in the skin (*31*).

1.3.2 Nodule palpation and Rapid Epidemiological Mapping of Onchocerciasis

With the massive expansion of control activities at the time the African Programme for Onchocerciasis Control (APOC) was set up (1995) (see section 1.6: *Onchocerciasis Intervention Programmes*), and the need to identify quickly areas at risk for high levels of onchocercal morbidity, a rapid and acceptable (non-invasive) diagnostic method was required. With this need in mind, the Rapid Epidemiological Mapping of Onchocerciasis (REMO) method was developed (*32*).

The REMO method is based on rapid epidemiological assessment of the prevalence of infection with palpable onchocercal nodules in a community, involving the examination of a sample of 30-50 adult men aged 20 years and above for onchocercomata. Mapping is started in a selection of villages perceived to be at high risk for onchocerciasis transmission (e.g. close to a major river in an area where biting blackflies are known to be present). For any village that proves at least mesoendemic (nodule prevalence ≥20%), a secondary village at least 10 km away is also surveyed.

The cut-off nodule prevalence of ≥20% corresponds roughly to a microfilarial prevalence of ≥40% (see also Coffeng et al. (*33*)) and is used as an indication of considerable risk for onchocercal blindness in a community. The levels of onchocerciasis endemicity were originally delineated by Prost (*34*) and adapted by the OCP as follows:

1. Hypoendemic: *mf* prevalence < 30%
2. Mesoendemic: *mf* prevalence ≥ 30% < 60%
3. Hyperendemic: *mf* prevalence ≥ 60%

1.3.3 Other diagnostic methods

Other diagnostic methods, that have applications when assessing the progress of control programmes, include the diethylcarbamazine (DEC) patch test based on the Mazzoti reaction (*35*)[35](*35, 36*),

serological assays for detection of parasite antigen (*37*) (dipstick assays (*38*)) and antibodies (antibody card tests (*39*)), in particular Ov16 for measurement of incidence in sentinel groups (*40, 41*), and rapid-PCR analysis of skin scratches and other tissues (*42, 43*). The use of metabolomics for the detection of presence of adult reproductive female worms is under investigation (*44*). McCarthy et al. (2012) have reviewed the current status of diagnostics for NTDs including onchocerciasis (*45*). Employment of these more sensitive methods of detection are becoming increasingly important as control programmes move towards elimination in order to quantify reductions of incidence in sentinel groups (untreated children under the age of 5 years) or to confirm infection-free areas.

1.4 Clinical manifestations of onchocerciasis

1.4.1 Generalised onchocerciasis

The generalized form of onchocerciasis is characterized by the presence of adult worms in onchocercomata and of millions of dermal *mf*. Individuals with generalized onchocerciasis often exhibit little skin inflammation and relatively mild skin atrophy and pigmentary changes, associated with this pronounced microfilaridermia. The main symptoms of generalized onchocerciasis, which represents roughly 98% of the cases (*46*), are pruritus (severe, often chronic, troublesome itching), a spectrum of dermal lesions (which can be exacerbated by pruritus), ocular lesions, and in areas of sustained and high transmission intensity, a high prevalence of blindness (*25, 47-51*).

By contrast, the rare, hyper-reactive form of onchocerciasis, also known as 'Sowda' (severe lichenified onchodermatitis), affects only a minority of individuals (*52*). These individuals are characterized by low microfilarial loads; low numbers of adult worms resident in large nodules (*53*); chronic, mostly unilateral hyper-reactive onchodermatitis, characterized by papular dermatitis; hyper-pigmented lichenified lesions; pruritus, and lymphadenitis (*54*).

1.4.2 Skin lesions and troublesome itch associated with onchocerciasis

There is an array of skin lesions resulting from onchocercal infection. These can range from mild, localised, papular rashes to chronic and extensively distributed lichenified, atrophied and depigmented skin. Murdoch et al. (1993) presents a thorough review (*2*). The skin is often covered in sores and pustules as a result of intense pruritic itching, which in itself can be chronic and lead to psychological and socio-economic problems, including social isolation, insomnia and absenteeism from work (*55-58*).

1.4.3 Ocular disease associated with onchocerciasis

Ocular lesions are common in infection with onchocerciasis. The first clinical signs of ocular onchocerciasis are usually the formation of small lesions on the cornea known as punctate keratitis, which are caused by the host's immune response to a combination of microfilarial and *Wolbachia* bacterial antigens released from dying microfilariae (*59-61*). (*Wolbachia pipientis* bacteria are endosymbionts of many filarial parasites, including *O. volvulus*; they are essential for the development, survival and reproduction of the parasite, and have a pro-inflammatory effect on the host that is deemed to be partly responsible for the clinical sequelae associated with onchocerciasis. (*62*))

These early lesions are reversible and can be attenuated by treatment with ivermectin. In the absence of treatment and with continued exposure, irreversible damage to the anterior and posterior chambers of eye may lead to progressive loss of visual acuity and narrowing of the visual fields, eventually leading to blindness (*63-65*).

1.4.4 Other disease manifestations

There is an association between *mf* prevalence and increased prevalence of epileptic seizures in populations with a high prevalence of onchocerciasis, in areas not co-endemic for neurocysticercosis due to infection with *Taenia solium* cysterci, or after controlling for this, known infectious cause of epilepsy (*66-68*).

The Nakalanga syndrome is a type of hypo-sexual dwarfism associated with *O. volvulus* infection, first described in 1950 and typified by stunted growth and sexual development, muscular atrophy, thin limbs and varying degrees of mental retardation (*69, 70*).

1.4.5 Excess mortality linked to parasite load

There is a statistically significant relationship linking excess mortality of the human host with increasing parasite loads (*4, 71, 72*), even when accounting for the excess mortality caused by blindness, which has previously been noted in large epidemiological surveys conducted amongst people with onchocerciasis (*73, 74*).

1.4.6 Differences in disease manifestation related to bioclime

There are striking differences in the morbidity experienced by human populations living in savannah versus forested areas (although as mentioned earlier, this is not a mere dichotomous divide). Whilst endemic communities living in forested regions may experience high levels of skin disease, there is often less ocular disease, even in communities where *mf* prevalence in the skin reaches nearly 100% (in those aged above 5 years) and the intensity of transmission (measured as the annual transmission potential or ATP) is greater than 50,000 L3 per person per year (*22, 75-77*).

The ATP in filarial infections is the equivalent of the entomological inoculation rate in malaria, but in the case of filariases, it measures not only the infective biting rate on humans, but also the average number of infective L3 stages that a person living in a particular area is typically exposed to in a year.

Conversely, communities living in savannah-type regions may experience relatively low values of ATP (typically up to 3,000-5,000 L3/person/year) whilst exhibiting high prevalence of infection with prevalence of blindness up to 30% (*78*). Experimental infection studies investigating the vector competence of various member species of *S. damnosum* sensu lato (s.l.) for *O. volvulus*, concluded that different and locally adapted parasite-vector combinations exist in these two bioclimes, such that the forest vectors are much less efficient at transmitting savannah-type parasites and vice-versa (*24, 79*).

Subsequent studies have shown genetic (and other) differences (including *Wolbachia* load (*80*)); between parasites collected in forest and savannah regions, which warrant consideration of these parasite populations as belonging to two different strains of the parasite (*81, 82*).

1.5 The *Simuliidae* vector hosts of onchocerciasis

Onchocerca volvulus is transmitted by *Simulium* species. The large majority of transmission in the African setting is due to members of the *S. damnosum* s.l. complex and the *S. neavei* group, with the former consisting of many related sibling species (i.e., species not easily distinguishable by morphological means, but by chromosomal methods leading to the identification of so-called cytospecies and cytotypes). The various member species of the *S. damnosum* complex are responsible for 95% of transmission worldwide (*10*). In the Latin American foci there have been 13 distinct transmission zones across 6 endemic countries, with many vector species and species complexes involved in transmission (*83, 84*). (Elimination of onchocerciasis has been achieved in some of these foci; e.g. in Colombia (*85*).)

Table A.1 in *Appendix A* lists the main vectors and their geographical distribution. In Africa the two different 'strains' of *O. volvulus* parasite (forest and savannah), are preferentially transmitted by vectors that are endemic to the bioclime of the parasite. *Mf* from persons infected with savanna-type parasites are less well adapted to develop into infective larvae within *S. damnosum* from the forest setting than *mf* from persons infected with forest-strain parasites, and vice-versa (*12*).

1.6 Onchocerciasis intervention programmes

Onchocerciasis was previously widespread in Sub-Saharan Africa, being endemic in at least 30 countries, while a relatively small number of cases was also seen in Latin America (6 countries) and the Yemen. However, the prevalence of infection and clinical manifestations in all affected regions have been reduced markedly due to the implementation of large-scale control programmes based on vector control and/or mass distribution of the microfilaricidal drug ivermectin.

1.6.1 The Onchocerciasis Control Programme in west Africa

The Onchocerciasis Control Programme in West Africa (OCP) was the first large-scale control programme, active from 1974-1975 to 2002 and covering 11 countries (Figure 1.2). The OCP started as a vector control programme, aiming to eliminate the disease by interrupting transmission for the duration of the adult worm lifespan. Large-scale vector control was achieved by weekly aerial spraying (of larvicidal insecticides) of *S. damnosum* sensu lato (s.l.) breeding sites to decrease aquatic stage populations, adult fly densities and thereby the transmission of infection. Initially the OCP covered 7 countries, but it was later expanded to prevent reinvasion of flies from neighbouring countries or uncontrolled areas within the original countries, covering 11 countries in total, namely Benin, Burkina Faso, Côte d'Ivoire, Ghana, Guinea, Guinea Bissau, Mali, Niger, Senegal, Sierra Leone and Togo (*86*).

When ivermectin (a drug previously for veterinary use) was licensed for humans and became available for treatment of onchocerciasis in 1987 (donated by Merck & Co.), control in OCP areas was supplemented with regular treatment of the at-risk human populations, and in some new areas this became the sole control strategy (*87*). After the closure of the OCP in 2002, onchocerciasis was considered largely eliminated in the OCP core area (an area including parts of the original 7 participating countries (*88*) and vector control was discontinued almost everywhere (with the exception of some Special Intervention Zones or SIZ (see below), where transmission was assessed to be ongoing despite more than 25 years of control; (*10*)).

After the closure of the OCP, the responsibility for mass treatment with ivermectin and surveillance of transmission has been devolved to national governments, which have implemented national onchocerciasis control programmes (now integrated within NTD programmes), often with APOC's technical assistance. The few areas in Togo, Benin, Ghana, Guinea and Sierra Leone, where control was not sufficiently successful, were designated as Special Intervention Zones (SIZ) (*89*) and control measures in these areas are ongoing as of 2014 (often using intensified ivermectin treatment).

Among its achievements, it has been estimated that the OCP was responsible for the prevention of 600,000 cases of blindness, making safe 25 million hectares of land for cultivation, and providing food production capacity for 17 million more people; additionally, 40 million people living in former OCP areas are now protected from the risk of onchocercal disease, and 18 million children had been born without the risk of becoming blind due to infection with *O. volvulus* (*90*).

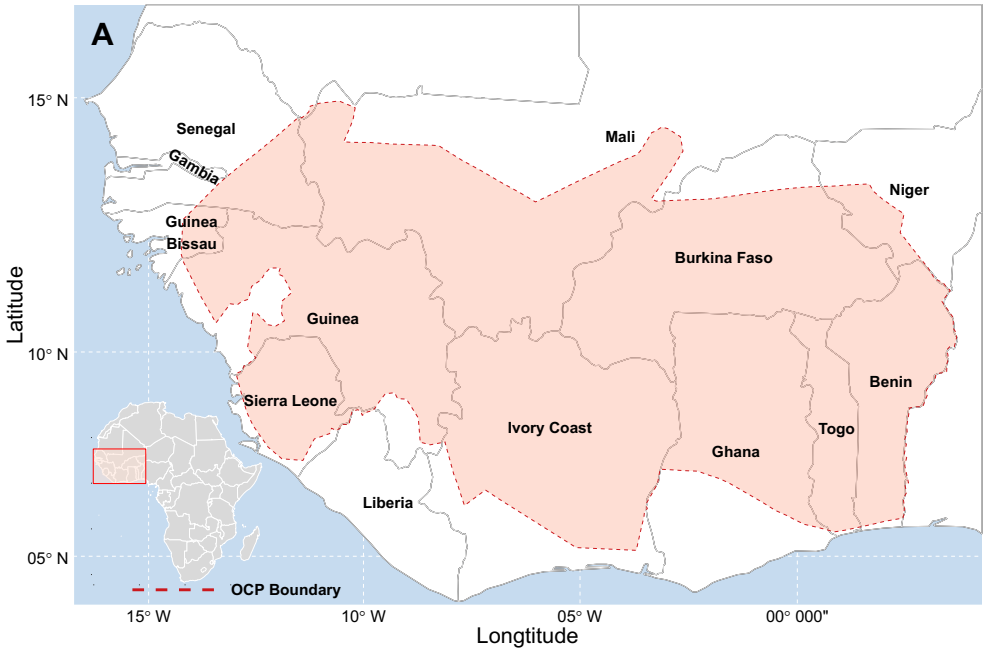


Figure 1-2: The area covered by the Onchocerciasis Control Programme (OCP).
The area enclosed by the dashed red line is the overall OCP area of operation including extensions to the original programme area.

1.6.2 The African Programme for Onchocerciasis Control

In 1995, the African Programme for Onchocerciasis Control (APOC) was created to organize national control programmes in 19 (West, Central, and East) African countries not covered by the OCP (*91*). Although there were some isolated vector elimination programmes within APOC (successful in selected foci of Equatorial Guinea (Bioko island), and Uganda but unsuccessful in Tanzania), APOC's main strategy was annual mass ivermectin treatment.

This strategy was organised around the philosophy of Community-directed Treatment with Ivermectin or CDTI, by which the communities themselves appoint distributors who are responsible for delivering and recording treatment on an annual basis. This philosophy was conceived to ensure the sustainability of prolonged treatment as a bottom-up approach that would empower the communities themselves (*92*). Previously, and in the OCP area, ivermectin was distributed by mobile teams, resulting in relatively low levels of geographic and therapeutic coverage. CDTI has been successful in this respect and has provided a platform onto which other interventions could be delivered (*93*). The

CDTI strategy has been adopted by all onchocerciasis endemic countries of Africa, including those under the former OCP.

Originally aiming to eliminate the infection as a public health problem, APOC set out to implement mass distribution of ivermectin treatment in all areas where onchocerciasis was considered an important public health problem (15 countries). APOC was initially replaced by the Programme for the Elimination of Neglected Diseases in Africa (PENDA) which was itself recently replaced by the Expanded Special Project for the Elimination of Neglected Tropical Diseases (ESPEN).

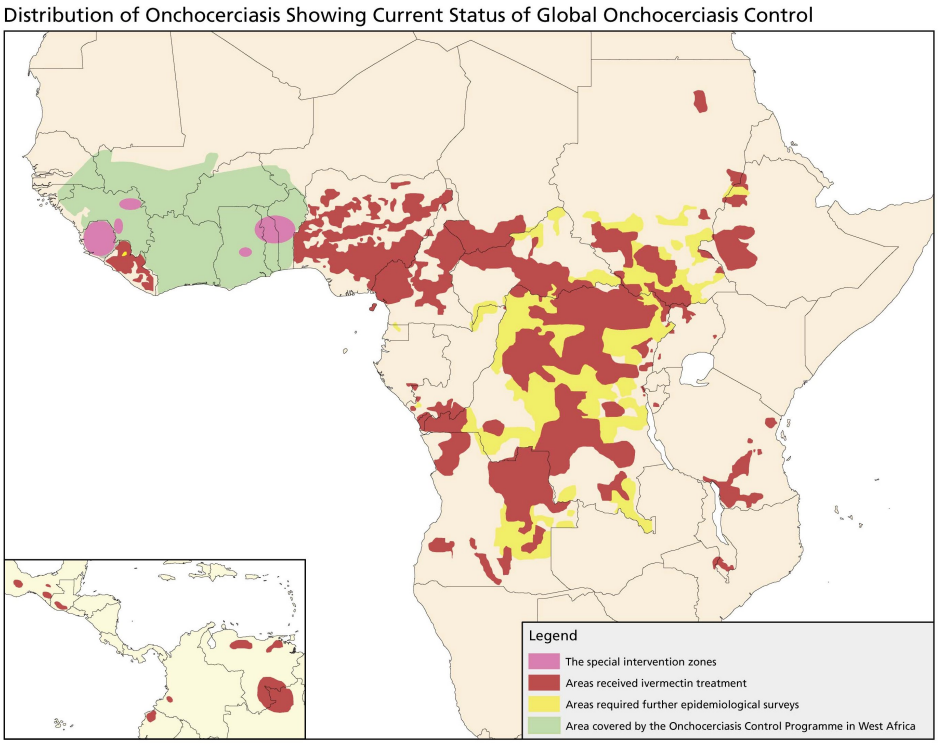


Figure 1-3: Global distribution of onchocerciasis.
The green area corresponds to the OCP area, and within it the purple ovals represent the special intervention zones (SIZ) that were set up in 2002. The areas in red are under mass distribution with ivermectin; in Latin America six-monthly by the Onchocerciasis Elimination Program for the Americas (OEPA) and in Africa annually by the African Programme for Onchocerciasis Control (APOC). Areas in yellow were under further evaluation at the time of preparation of this map in 2006 (from Basáñez et al. (2006) (*10*)).

1.6.3 Other Intervention Programmes

Mass ivermectin treatment is also on-going in more circumscribed foci in the Yemen and Latin America, but not under the modality of CDTI; instead, mobile teams distribute the drug. The Onchocerciasis Elimination Program for the Americas (OEPA) aimed from its onset at elimination of the infection by 6-monthly mass distribution of ivermectin (*94, 95*). This, more frequent strategy, was considered feasible because of the highly focal nature of onchocerciasis in Latin America, the small populations at risk, the relatively good health infrastructure available, health capacity and public funding, and the relatively lower vector competence for *O. volvulus* of some simuliid species in Latin America.

This strategy has led to interruption of transmission and possibly parasite elimination in some circumscribed foci (*95, 96*), including the Santa Rosa focus in Guatemala and the Lopez de Micay focus in Colombia. Other foci in Mexico, Guatemala, Ecuador, and northern Venezuela are in the process of applying for certification of elimination. In some foci (in Mexico and in the Amazonian focus straddling Venezuela and Brazil), a strategy of 4 ivermectin rounds per year has been implemented (*97*).

1.7 Intervention Strategies

1.7.1 Larviciding of vector breeding sites

The initial strategy employed to eliminate onchocerciasis as a disease of public health importance was weekly aerial larviciding of vector breeding habitats (chiefly with organophosphates). Operations were particularly focused against eliminating the blinding strain of the parasite and larviciding operations commenced in stretches of river closest to villages with high incidences of blindness (*98*). The OCP began operations in January 1974, and the first larviciding treatments began in February 1975. A roster of 7 different larvicides was in use at any one time (*99*). The larvicide used in a

particular river stretch was chosen on a weekly basis, after consideration of: a) the hydrologic conditions (how quickly will the larvicide be diluted down / washed away, below the effective larvicidal threshold); b) the environmental impact (unintended negative impacts on non-target organisms, e.g. aquatic fauna), and c) how best to minimize the possible risk of emergence of resistant simuliid vector populations. An extensive aquatic monitoring programme was put in place prior to control operations to evaluate and manage the impact of long-term use of insecticides in affected riverine ecological systems (*100*).

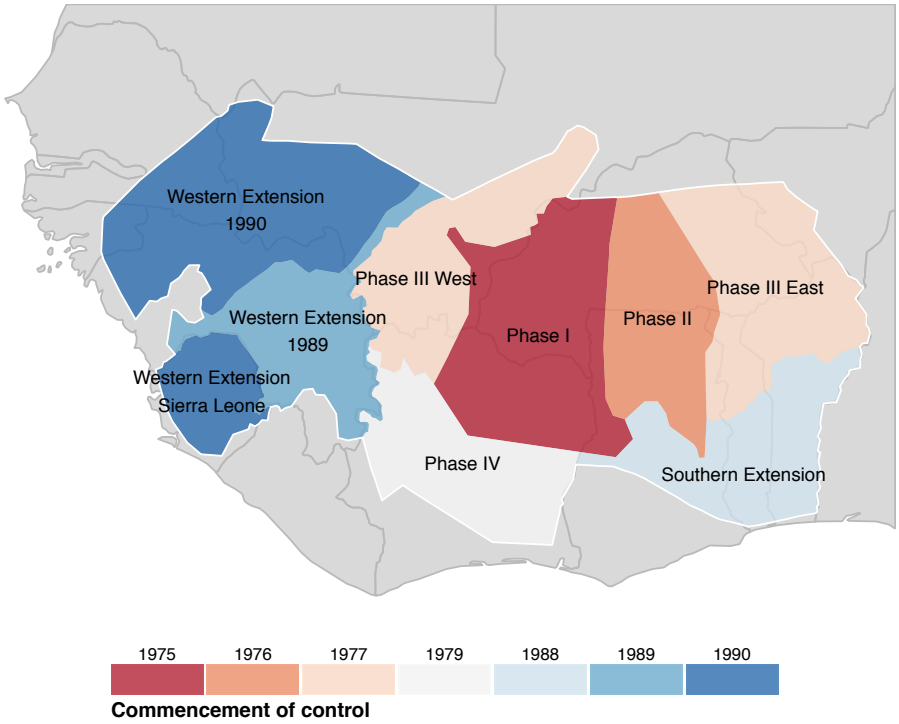


Figure 1-4: Operational phases of the OCP.
Larviciding operations were carried out starting in February 1975 in Phase I. Phases were added chronologically. In 1977 the decision was made to expand the programme from its current limits in 7 countries (Phases I, II, III East and III West) to incorporate infected river basins in neighbouring countries, covering a total area of more than 1.25 million Km^2.

At the end of 1977, 645,000 km^2 of onchocerciasis-infested land were covered by larviciding operations (Phases I, II, III-East and III-West on Figure 1.4). Despite great progress made in these initial core areas, achieving the control aims was hampered by re-invasion of infected flies from river basins in the areas surrounding these controlled zones. Further extensions were set-up to protect the

original area and to extend the successful strategy in areas not previously covered (Phases IV, Southern Extension and the 1989 Western Extension) (*99*).

1.7.2 Mass treatment with ivermectin

Ivermectin (a macrocyclic lactone derived from fermentation products of *Streptomyces avermectilis*) had been developed for treatment of veterinary parasitic nematodes mainly affecting ruminants of economic importance. In 1987, it was licensed for use on humans and it was shown to be highly efficacious against the microfilarial stage of *O. volvulus* (*101*). The availability of the drug ivermectin (Mectizan™) changed the strategy of onchocerciasis control from antivectorial to antiparasitic measures or a combination of both (*94, 95*). Merck & Co., who manufacture ivermectin, have committed to donate the drug as long as necessary to onchocerciasis control and elimination programmes worldwide through their 'Mectizan Donation Program' (MDP), now part of the Global Health Task Force with headquarters in Atlanta, GA, USA.

1.7.3 Mode of action of ivermectin

Ivermectin is a highly effective microfilaricide; however, it has a limited efficacy against the adult stage of *O. volvulus*. The *mf* in the dermis of a human host are significantly cleared within 72 hours after taking the drug and a 95-99% reduction is achieved within 1-2 months (*102, 103*), preventing onward transmission via blackfly vectors for a number of months, and abating the onset of further morbidity caused by *mf* load in infected persons. However, microfilarial loads recuperate somewhat between two consecutive treatments (due to a combination of skin repopulation by newly produced *mf* by established worms and the production of *mf* by new worms acquired through reinfection).

There is also a reduction in fecundity of adult female worms (embryostatic effect) (*103*), slowing the repopulation of a host's skin by microfilariae. Of current debate is whether this reduction in fecundity

is a temporary but reversible reduction (*104*), or if there is a cumulative and permanent reduction in fecundity with subsequent treatment rounds (*105, 106*). Elucidation of this difference has clear implications for control and elimination programmes and mathematical modelling studies have provided projections of the impact of long-term ivermectin distribution based on these different assumptions, accompanied by health economic evaluations (*107-109*).

1.7.4 Sub-optimal responses to repeated treatment with ivermectin

Following a single dose of ivermectin, a patient's skin may recover to 6%, 17% and 40% of its baseline *mf* load at 6, 12 and 24 months respectively, as the female worms recover their pre-treatment fecundity rate (*103*). After more than 25 years of repeated doses of ivermectin across the former OCP area, in some West African countries and some APOC project areas, there is concern that the issue of potentially emerging ivermectin resistance must be addressed (*110*). Faster than expected rates of repopulation of *mf* in the skin have been reported, primarily in areas of Ghana with a long history of control with ivermectin (since *circa* 1987) (*111-113*). It is hypothesised that this could be due to emerging resistance to ivermectin's effect on fecundity of the adult female worm (reduced embryostatic effect) (*113*) or due to other (non-resistance) causes (*114, 115*) such as poor programmatic performance, low treatment coverage and compliance, and intense transmission.

1.8 Mapping of human onchocerciasis

1.8.1 Initial attempts to map endemic foci of onchocerciasis

The World Health Organization (WHO) Expert Committee on Onchocerciasis was first convened in 1953, and from their first report in 1954 (*116*), following a review of available epidemiological surveys at the time, knowledge of disease distribution was poor:

"For many regions where the infection is known to exist, however, data are not yet sufficient to evaluate the real social importance of the disease, and more detailed surveys are necessary." (WHO, 1954, p. 6)(116)

In 1960, under the auspices of ORSTROM (Office de la Recherche Scientifique et Technique Outre Mer), the then newly created '*Organisation de Coopération et de Coordination pour la Lutte contre les Grandes Endémies'* (l'OCCGE) launched the first anti-vectorial campaign against *Simulium damnosum* in the upper stretches of the Black Volta (a region spanning the Niger / Burkina Faso borders in the southern regions of these countries) (*117*). This campaign was informed and directed by knowledge of vector breeding sites from entomological experts working in the field. McMahon (1967) offers a review of control activities and against *Simulium spp*. up to that time (*118*).

1.8.2 The inception of the OCP and improving knowledge about disease distribution

The initial campaign was subsequently followed by a much larger operation conducted from 1966–1974, which also included parts of northern Côte d'Ivoire, and covered 50,000km^2. The results from this successful campaign were included in the report of the Preparatory Assistance Mission to the Governments of 6 West African countries, and gave rise to the OCP. The second report of the WHO's Expert Committee on Onchocerciasis (1966) highlighted the knowledge requirements for such an operation to be successful, namely, extensive entomological surveys in both wet and dry seasons, extending beyond the control area, involving international co-operation where necessary (*119*).

Huge amounts of data were collected which enabled highly targeted larviciding of river courses infected with *S. damnosum (120)*. Figure 1 from page 7 of this report is reproduced as **Figure B.1** in **Appendix B**; *Historical Maps of Onchocerciasis Prevalence in West Africa*. It displays the state of

knowledge in 1966 about human onchocerciasis distribution, based on entomological and parasitological village surveys.

Despite substantial progress, there were still large gaps in knowledge about the distribution and severity of endemic foci. In addition, there was no knowledge about *how endemic* these foci truly were, which is essential to target control efforts. Up until this time, and as operations were largely focussed against the vector, there was little knowledge about the numbers of people infected and blind. It is mentioned in the 1966 report that there were an estimated 350,000 infected persons in Nigeria, of whom 20,000 would be blind. In the 1987 report this figure was reported as 6.96 million and 113,600 blind (*119*), and by 1995, with more detailed surveys this figure was revised to 3.3 million infected with 100,000 blind (*7*).

In 1987, the WHO published its Third Report from the Expert Committee on Onchocerciasis (*119*). To the best of the candidate's knowledge it contained, for the first time, estimates of the numbers at risk, infected and blind on a per country basis for the entire African region. However these data were again based on small sample surveys, extrapolated to the country level:

> *"In many countries the total figures for the prevalence of infection and blindness have been estimated by extrapolation from relatively small sample surveys and without an accurate knowledge of the local distribution of endemic foci." (WHO, 1987, p. 9)(119)*

1.8.3 Mapping disease distribution in the Western Extension of the OCP

With the extension of the OCP to the West a more sophisticated and detailed attempt was made to map endemic foci and areas at high risk of onchocercal blindness (*121*). The Western Extension was divided in zones with highly similar entomological and ecological characteristics. An estimate of the

population at risk was made for each zone. Within each zone there were a number of defined river stretches. For each river stretch, 2 first-line villages (villages within 5km of a riverbank and with no other villages between them and the river bank) were randomly selected and then 2 other non-first-line villages were randomly selected for parasitological surveys. The data from these surveys were aggregated for each zone and then some estimates of the number infected and areas at risk of infection were made and used for directing control operations in the Western Extension.

This represented a substantial step forwards in the methodology used for mapping onchocerciasis distribution, but it still did not make use of any spatial autocorrelation between closely situated villages, and only a small fraction of the population could be sampled. In the WHO's Expert Committee on Onchocerciasis' Third Report, the estimate for the total number of people infected at the time (1995) was put at 17.7 million. This was based on knowledge collected over some years previously (*7*).

1.8.4 Mapping disease distribution in APOC

When APOC was initiated in 1995, the method of control was only based on ivermectin distribution, and therefore much more detailed maps of the population at risk and infected were required for this purpose for which REMO was developed (see Section 1.3.2 *Nodule Palpation And Rapid Epidemiological Mapping Of Onchocerciasis* for a detailed description of REMO). The prevalence data obtained during the surveys were plotted and zones of endemicity levels delineated. Unfortunately, the method used to delineate zones is not described (*32, 122*), but seems to imply the use of expert knowledge from entomologists, epidemiologists and geographers. Subsequently, maps using kriging techniques were developed from the REMO survey results, but these maps and the methods used to generate them are unpublished and only available internally to APOC (*Dr Wilma Stolk, pers. comm.).*

1.8.5 The importance of accurate maps of disease distribution

Given the aims and intervention strategy of APOC—namely to eliminate the public health burden of onchocerciasis in Africa and to distribute ivermectin annually to the communities at higher risk of morbidity—for the first time it became vital to delineate the level of endemicity for the purposes of planning the control programme. The original goal of the programme was to eliminate onchocerciasis as a *disease of public health importance*. To achieve high geographical coverage in those areas most at need, with the resources available, the decision was taken to treat areas that were mesoendemic or hyperendemic only (*123*) as the risk of developing severe and disabling morbidity is directly linked to intensity of infection (*25*). The relatively benign morbidity found in hypoendemic areas was considered less of a public health burden so these areas were not targeted (as APOC was not then aiming at elimination).

According to the last complete REMO survey results (2005), more than 37 million people were infected with *O. volvulus* (doubling previous estimates) and a further 120 million people live in areas at risk of infection (*7, 124*). With more recent progress in the coverage of REMO surveys, knowledge about the distribution of infection within APOC's intervention area has been refined, calling for an updated estimate of the numbers of people suffering from onchocercal infection and disease in the APOC area (*125*).

1.8.6 Recent efforts to map onchocerciasis and other Neglected Tropical Diseases

In the last decade there has been a substantial and renewed focus on the re-branded and now called Neglected Tropical Diseases (NTDs) (*126-129*). These are tropical diseases that have been long neglected by political, economic, industry (R&D) and scientific communities, but that affect the bottom billion of the world's population, accentuating and perpetuating a cycle of poverty. Vital to the effort in bringing NTDs to the fore, has been the preparation of accurate disease distribution maps for the targeting of control programme resources. The production of continuous prevalence maps has

made increasing use of model-based geostatistics (MBG) (*130*), a branch of statistics that combines spatial analysis of ground-truth data with knowledge of environmental determinants to try and predict disease risk.

MBG often employs Bayesian inference (B-MBG) for spatial prediction and robust characterisation of uncertainty surrounding those predictions (*131*). Detailed prevalence maps have been presented for a variety of directly- and indirectly-transmitted human helminthiases, e.g. (*132-134*), and vector-borne diseases, e.g. (*131, 135, 136*), including lymphatic filariasis (*137*), and loiasis (*138*). Loiasis is of particular relevance to onchocerciasis due to the risk of severe adverse reactions to ivermectin in people co-infected with *Loa loa* (another filarial parasitic nematode) and onchocerciasis. Recently a spatial risk map for onchocerciasis in Ghana and Burundi was published, developed using a spatial generalised linear mixed model (GLMM) (*139*). However, no estimates of infection prevalence were made, and the published data used in this case study were far more limited than those data available from the OCP epidemiological dataset (*139*). As a result, no continuous disease distribution maps yet exist for onchocerciasis in West Africa.

1.8.7 Historical maps of onchocerciasis prevalence in Africa

Appendix B reproduces several historical maps from WHO reports and other publications for the purposes of illustrating changing knowledge of the distribution of onchocerciasis.

1.9 Thesis aims and objectives

Among the main determinants of the duration of ivermectin-based control and feasibility of elimination is the initial, pre-control level of endemicity, measured in terms of the community prevalence of infection with O. volvulus microfilariae (*109, 140*). Microfilarial prevalence is positively correlated with transmission intensity (*141, 142*), measured in terms of the annual

transmission potential (ATP)—the yearly number of infective larvae of O. volvulus potentially received by a person maximally exposed to simuliid bites in the same community (*141-143*). Therefore, and to further understanding of the feasibility of onchocerciasis elimination, this thesis embarked on the task of mapping the distribution of the baseline microfilarial prevalence in the former OCP area.

Recent modelling studies (*109*), have indicated that reaching the operational elimination thresholds suggested by APOC (*108*) may be feasible, with an annual ivermectin distribution strategy, in areas with a baseline microfilarial prevalence less than 60% (the cut-off for hyperendemicity), provided high levels of therapeutic coverage and treatment compliance are maintained. By contrast, in initially highly hyperendemic areas (prevalence of 80% or above), biannual ivermectin distribution or deployment of complementary strategies (e.g. focal vector control if possible; delivery of anti-Wolbachia therapies (*144*)) may be necessary.

Mapping of infection prevalence is of importance not only for targeting current or novel control efforts appropriately, but also, once local elimination has been achieved, or is deemed possible to achieve within proposed periods, to predict areas at risk of recrudescence from geographically contiguous areas still harbouring infection. Finally, mapping the baseline prevalence can help to determine the population at risk and infected prior to the inception of control interventions, essential denominators to calculate the initial burden of disease and its changes as a result of such interventions (*145*). These aims will be achieved through the presentation of the following chapters.

1.9.1 Chapter 2: Investigating and modelling spatial autocorrelation in Onchocerciasis Control Programme data

The aim of Chapter 2 is to investigate and identify spatial autocorrelation in the survey data from the Onchocerciasis Control Programme in West Africa. The pre-control data will then be used to construct a map of pre-control prevalence of microfilarial prevalence using several different methods,

with the aim of validating the best method to estimate prevalence in unsampled areas and whether the incorporation of spatial information improves predictive accuracy.

1.9.2 Chapter 3: Bayesian geostatistical mapping of the distribution of human onchocerciasis in west Africa prior to control interventions

The aim of Chapter 3 is to improve and refine estimates of pre-control prevalence from Chapter 2, through the incorporation of environmental covariates in the model building process for the objective of accurately estimating the infected and at-risk pre-control population. Additional information from environmental covariates will be considered in a step-wise fashion, aiming to fit the best explanatory model possible, that also incorporates a spatial covariance model with the aim of producing the most accurate estimates of prevalence in unsampled areas. An additional aim is to characterise uncertainty in the estimates and this will be achieved through use of Bayesian methods.

1.9.3 Chapter 4: Estimating numbers infected and the burden of disease caused by onchocerciasis prior to control interventions

The aim of Chapter 4 is to use the output maps from Chapter 3 to estimate the burden of disease caused by onchocerciasis through the causes of low vision, blindness and skin lesions. These maps will be combined with population distribution maps to re-estimate the number of persons infected and at-risk pre-control. A comparison will be made between the different disability weights arising from different version of the Global Burden of Disease study, and an assessment of the DALYs saved through the intervention of the Onchocerciasis Control Program will be made by using updated population count maps.

1.9.4 Chapter 5. Investigation of socioeconomic factors affecting compliance in Community-Directed Treatment with Ivermectin programmes

Previous chapters have focussed on quantification of onchocercal disease infection and burden. In this chapter, the focus is on qualitative research to investigate socio-economic factors that may influence the performance of control programmes and possible factors affecting a communities' response to ivermectin treatment. In particular, this study focuses on two regions of Ghana where there has been in excess of 25 years of control with ivermectin, yet where some communities appear to have a sub-optimal response to treatment (a higher than expected prevalence of *mf* given the length of treatment and initial endemicity level).

An analysis of an extensive household survey carried out by the candidate during fieldwork in 280 households at 13 communities within the Northern and Brong-Ahafo regions in Ghana is undertaken to find risk factors related to self-reported onchocerciasis status and treatment compliance with community-directed treatment with ivermectin at the individual, household and community level.

1.9.5 Chapter 6. Final discussion and future directions of work

The aim of this chapter is to summarise what has been presented and to outline directions for future work and research needs arising from this thesis.

Chapter 2. Investigating and modelling spatial autocorrelation in Onchocerciasis Control Programme data

2.1 Overall aims and objectives

The overall aim of this chapter is to develop a continuous surface of pre-control microfilarial prevalence using pre-control data from the OCP epidemiological dataset and kriging methods. This will be accomplished through the following broad objectives: a) to define and select pre-control survey data from the OCP epidemiological dataset; b) to identify and characterise any spatial autocorrelation present in the selected data; c) to develop surfaces of microfilarial prevalence using these data, and d) to identify the best method of generating a surface of microfilarial prevalence and assessing if the inclusion of explanatory covariates can improve surface estimates.

2.2 Introduction

In addition to understanding the distribution of onchocerciasis endemicity for the purposes described in Chapter 1 (see **Section 1.9** *Thesis Aims and Objectives*), mapping the baseline infection prevalence can also help to determine more accurately the population at risk and infected prior to the inception of control interventions, which are essential denominators to calculate the initial burden of disease and its changes as a result of such interventions. The Global Burden of Disease study (GBD), is an incidence-based project that aims to estimate the total amount of disability caused by disease, injuries and risk factors (see Chapter 4, **Section 4.1.2**; *The Global Burden of Diseases, Injuries, and Risk Factors Study*).

Due to the incidence-based approach taken by the GBD study, and the way in which control interventions against onchocerciasis have shifted the endemic equilibria—resulting from those regulatory processes that exist between the parasite-host, parasite-vector, and host-vector interactions in the absence of control perturbations—a "gold standard" estimate of the initial population infected is required, from which estimates of incidence of onchocerciasis may be estimated.

Although such estimations have been conducted for the projects under APOC's jurisdiction (*125*), no such values truly exist for the former OCP area, in which previous figures are acknowledged to have been grossly underestimated (*7, 116, 119, 120, 146*).

According to the latest rapid epidemiological mapping of onchocerciasis (REMO) surveys, more than 37 million people would have been infected with O. volvulus (doubling previous estimates) and a further 120 million people live in areas at risk of infection (*124*). With progress in the coverage of REMO surveys, new areas of endemic infection are being uncovered within the APOC intervention area, calling for an updated estimate of the numbers of people suffering from onchocerciasis and its disease sequelae.

In addition, previous rounds of the GBD assessment of onchocerciasis did not have access to empirical estimates of the population infected in former OCP areas in West Africa, with figures for the population at risk and infected based on the limited data available at the time. It was, therefore, opportune to make a complete re-estimate of the burden of disease incorporating better estimates of the pre-control distribution of onchocerciasis, using the pre-control survey data available from the OCP epidemiological database.

2.3 Methods

2.3.1 Parasitological survey methods and the use of crude prevalence data

The methods used in the epidemiological surveys of the OCP have been previously described (*147*). At each survey a complete census of the village was conducted, and approximately 84% of persons enumerated in the census were examined (*148*). Parasitological assessment comprised taking bloodless skin snips (with a 2 mm Holth-type corneoscleral punch), one from the right and one from the left iliac crests. Biopsies taken with this punch are relatively similar in size, weighing between 1 and 3 mg (*149*). The skin snips were placed in distilled water for 30 minutes and examined by

microscopy for presence (and number) of emerged microfilariae. Negative snips were re-incubated for up to 24 hours in saline solution and examined again.

Reported prevalence at the village-level was age- and sex-standardised to the OCP reference population (*29*). Due to the invasive nature of the biopsy procedure and relatively low prevalence in persons under 5yrs old (although this depends on the force of infection), skin snipping was performed in the population, aged 5yrs and above only. Standardised prevalence data were available in the database, but age- and sex-adjusted counts of positive persons were not. It would have been possible to back-calculate these figures, but although standardised prevalence was provided by sex, the data by age-group were not available. The linear correlation between the standardised and unadjusted prevalence was calculated and found to be highly linear, with a high determination coefficient (R2 > 0.95), therefore, the unadjusted count data was used for this study.

2.3.2 Description of the Onchocerciasis Control Programme epidemiological database

Relevant survey data were extracted from the OCP epidemiological database, which contains 5,816 surveys conducted in 2,581 (geo-referenced) villages across all operational phases of the OCP in the 11 participating countries in West Africa. **Figure 2-1** plots the locations of all village sites in the OCP epidemiological database. There were 9 operational phases (summarised in **Table 2.1** and depicted in Chapter 1, **Figure 1-4**). A total of 737 villages had parasitological surveys that met the inclusion (pre-control) criteria. **Figure 2-2** plots the location of the villages selected to be representative of the pre-control situation in the OCP area (see also **Section 2.3.3** *Selection of Pre-Control Data* and **Figure 2-3**).

The endemicity level of the selected villages are shown in **Figure 2-3**, with the endemicity classes defined as non-endemic/sporadic endemicity: < 10% microfilarial prevalence; hypoendemic onchocerciasis: ≥ 10% and < 35%; mesoendemic: ≥ 35% and < 60%; hyperendemic: ≥ 60% and < 80%, and highly hyperendemic (or holoendemic): ≥ 80% prevalence. Although in the OCP it

was standard epidemiological practice to adjust microfilarial prevalence by age and sex according to the OCP reference population (*29*), in this study crude microfilarial prevalence was used because counts of persons examined were not available in the data by age cohort.

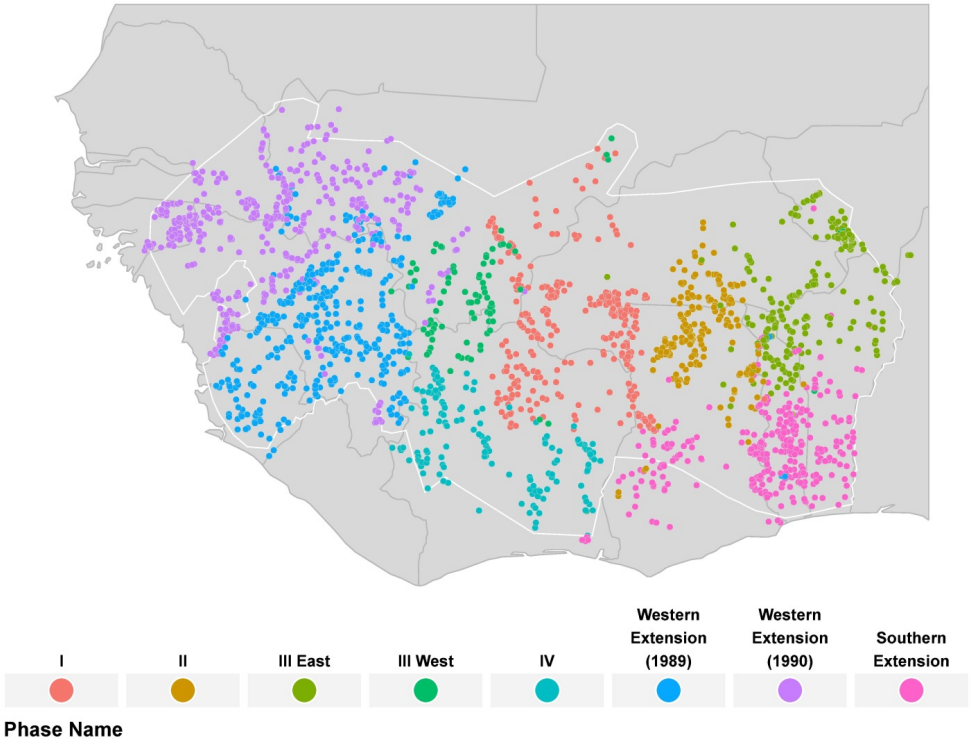


Figure 2-1: Location of villages in the OCP epidemiological database.
This map shows the position of all villages extracted from the OCP dataset with valid geo-referencing data. The data are coloured according to the operational phase that is listed for that village in the database (see Figure 1-4). Some villages have either an incorrect phase code or an incorrect geo-reference due to the interspersed colour patterns visible.

In 331 locations, where 393 parasitological surveys had been carried out, the data were missing geo-referencing information or were coded with erroneous geo-references (208 locations were coded with X/Y coordinates of -1/-1 or 0/0). A further 175 locations had missing data for numbers of persons examined and/or numbers of persons positive at each site (data coded as '-1'). After trimming these data, 5,248 surveys at 2,224 unique locations within the OCP remained. The surveys were divided between phases which received, primarily, mass treatment with ivermectin and those where vector control was the primary method of intervention between 1975 and ~ 1990. There were 2,525 surveys in chemotherapeutic OCP phases and 2,723 surveys in vector larviciding phases.

2.3.3 Selection of pre-control data

Each operational phase of the OCP (**Figure 1-4**) is comprised of several contiguous river basins or watersheds. The OCP database contains data on the start date of control within each river basin, as well as the overall start date for that operational phase (usually the same as the earliest date of control in any river basin within that phase). However, the initiation of control was not implemented concomitantly across the river basins in a single phase. The size of a river basin referred to in the OCP dataset may be quite small and sometimes refers to a small local tributary, which is part of a larger contiguous hydrological watershed. Therefore, control can be quite localised. Some river basins did not receive control until some years after the start of control in other river basins within that phase, e.g. in Phase I the listed start of control ranged from February 1975 to March 1979. Therefore, the start date of control of the river basin in which a village was located was used to define cut-off values for selecting surveys as pre-control.

In the phases that started with vector larviciding (Phases I, II, III-E, III-W and IV) the criteria that all surveys must have been carried out before the start of control were relaxed somewhat. Due to the method of control prior to mass treatment with ivermectin (which affects the microfilarial stages of the parasite) being vector larviciding, decreases in microfilarial prevalence would have been quite slow at first, especially in hyperendemic areas (*88*). For this reason, in all phases where vector control was the primary method of control, surveys were allowed that were conducted within 3 years of the start date of control for that river basin to be chosen (according to epidemiological trends in the OCP core area presented by Hougard et al. (2001) (*88*)). This increased the pool of surveys from which to select in these areas, from 209 to 443. Finally, at locations where there was more than one survey available, the earliest of the surveys was taken to be the most representative of the pre-control endemic equilibrium. After removal of 107 later surveys, 336 pre-control data points remained in these phases.

In phases that commenced with mass ivermectin treatment (the Western Extension and Southern Extension Phases), selection was restricted to surveys conducted prior to the start date of control

within river basins. Ivermectin is an effective microfilaricide (*101*) but microfilarial load responds to the drug in a dynamic fashion (due to the combined temporal dynamics of the microfilaricidal and embryostatic effects) (*103*). Although the very nonlinear relationship between microfilarial load and prevalence (*142*) would suggest that average reductions in prevalence occur more slowly than those in infection intensity, uncertainty in the times of surveys in relation to the times of treatment indicated that it was best only to use endemic equilibrium prevalence.

There were 453 surveys eligible at sites where multiple parasitological surveys were available. The earliest surveys were selected as the ones that would be most representative of the pre-control equilibria. This left 401 surveys available in these phases and a total of 737 data points overall (**Table 2.1**).

Phase Name		**Start Of Control**	**Type Of Intervention**	**# Of Pre-Control Surveys (n = 737)**
I		February 1975	Larviciding	115
II		January 1976	Larviciding	61
III East		March 1977	Larviciding	87
III West		February 1977	Larviciding	27
IV		March 1979	Larviciding	47
Southern Extension		February 1988	Larviciding	83
Western Extension	1989	March 1989	Ivermectin	317
	1990	April 1990	Ivermectin	
	Sierra Leone	March 1990	Larviciding + Ivermectin	

Table 2-1: Distribution of pre-control points across OCP phases.
Table displays the number of pre-control input data points in each operational phase of the Onchocerciasis Control Programme in West Africa. The earliest commencement of control operations within each phase is listed as well as the initial or main type of intervention.

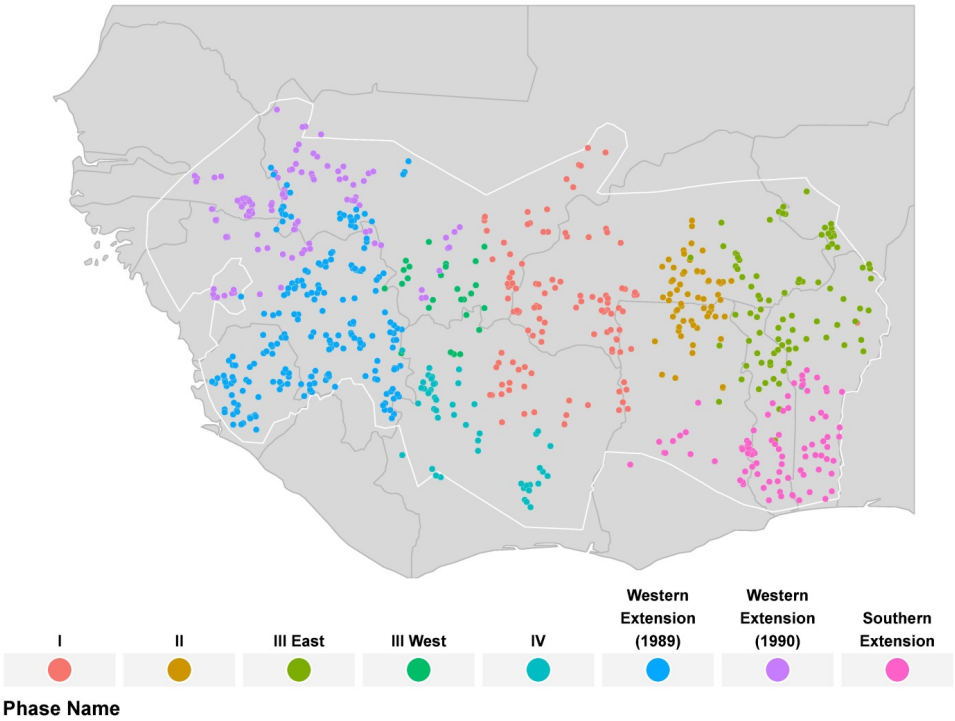


Figure 2-2: Location of pre-control villages selected for use in this study.
The data points (n=737) are stratified by operational phase of the OCP. The overall spatial distribution of the selected points displays good agreement with the overall distribution of the full dataset.

After data cleaning, errors due to miscoded phases or erroneous coordinates were reduced as can be seen by comparing **Figure 2-1** (all locations) versus **Figure 2-2** (pre-control villages only). The selected villages are stratified within phases with very little incorrect phase coding or geo-referencing information evident (overlapping colours in the plotted points) between phases except between two of the Western Extension phases. However, these phases started control at the same time, via the same method so this was not thought to be a significant confounding issue. The distribution of pre-control points shows good congruence with the overall distribution of all data points in the database. **Figure 2-3** displays the 737 selected communities, coloured by endemicity level.

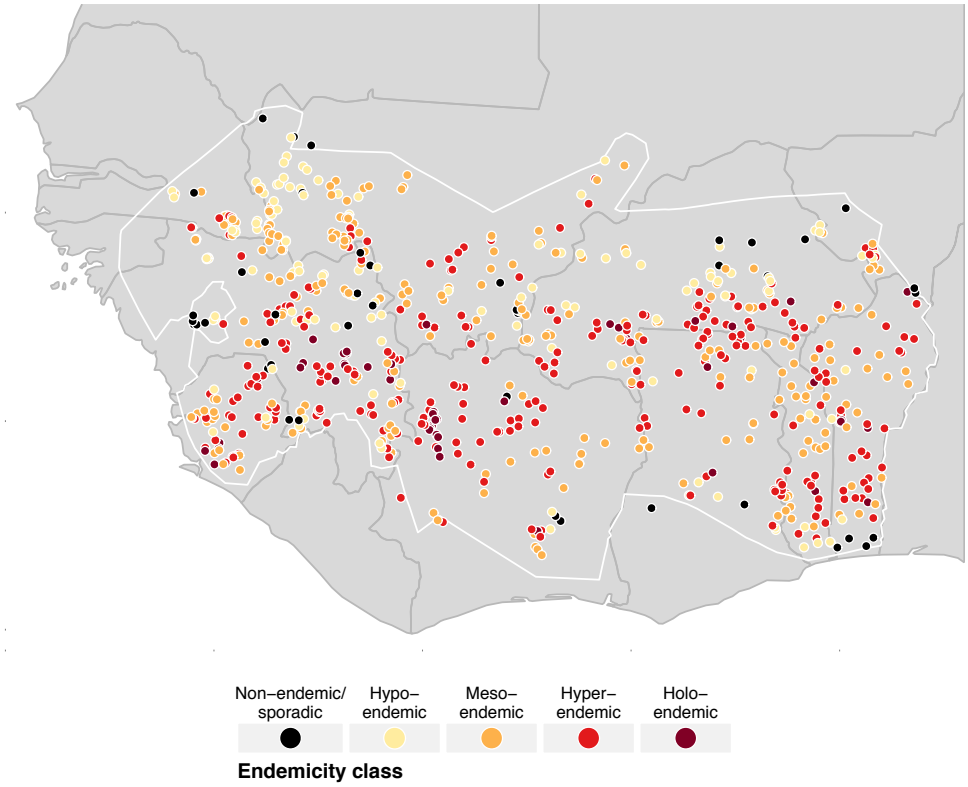


Figure 2-3: Endemicity class of pre-control villages.
This maps displays the distribution and prevalence of the selected pre-control villages. Two additional endemicity classes have been defined for this map in addition to those outlined in **Section 1.3.2** *Nodule palpation and Rapid Epidemiological Mapping of Onchocerciasis* to highlight areas of particularly low or high epidemiological importance. These are *Non-endemic/sporadic* (<10%) and *Holo-endemic (>80%)*.

2.3.4 Onchocerciasis risk factors

There are a number of risk factors thought to be associated with onchocerciasis infection. Distance to a vector breeding site and the density and composition of blackfly at one of their breeding sites are known to be important determinants of onchocercal infection. In order to make better predictions of prevalence in un-sampled areas, three environmental covariates were derived that could be used as de-trending (explanatory) covariates in interpolative and kriging models of onchocerciasis, microfilarial prevalence.

2.3.5 Distance to river and hydrological rank

Blackfly vectors of the Simulium damnosum s.l. species complex predominantly breed in fast flowing rivers and streams that satisfy certain ecological conditions such as being highly oxygenated, with larvae surviving in rivers with discharge volumes in the range of 80-150 cubic metres per second with a current velocity of 0.5 to 2.0 metres per second (*150*). Blackfly biting rates on humans decrease rapidly at increasing distances from a riverine breeding site (*151*). Whilst the OCP located and surveyed many vector breeding sites, there is incomplete information on the location and density of blackflies at breeding sites across the whole region. Therefore, distance to the nearest perennial river stream was used as a proxy for this risk factor.

The AQUASTAT dataset compiled by the Food and Agriculture Organization of the United Nations (FAO) contains a GIS shapefile of all rivers in the FAO Atlas of Resources and Irrigation in Africa (*152*). The map was digitised in 1994 and contains 32,636 features (stream segments), which are classified according to whether they are perennial or intermittent, and also by hydrological rank of each river segment.

To help account for the difference in density of blackflies at a breeding site, the hydrological rank of the nearest perennial stream segment was used as a proxy, with higher hydrological rank relating to a presumed higher density of blackfly at a vector breeding site. Hydrological rank is classified according to Strahler's stream order system (*153*), which proposes that a river or stream segment be given a rank according to the number of upstream tributaries it possesses. Therefore, a first order stream segment has no upstream tributaries and is known as a headstream. As rank increases so does the size of a river (as more tributaries feed into it), which can give an indication of the magnitude of the river in that location compared to the upstream area.

The map of rivers was cropped to the OCP area and the Euclidean distance to the nearest perennial river segment was calculated for each of the 737 pre-control villages (**Figure 2-4A** across study area and **Figure 2-4B** for zoom detail).

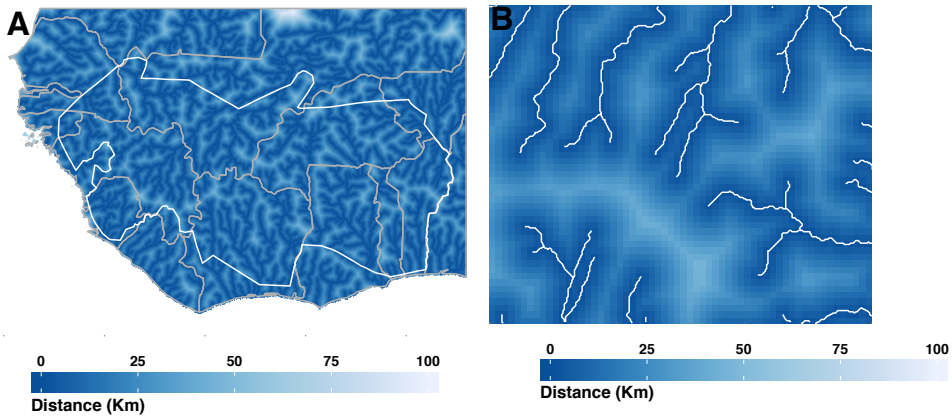


Figure 2-4: Euclidean distance to nearest river.
Panel A) Each 25 km^2 pixel is coloured according to the Euclidean distance from the pixel centre to the nearest stream segment from the Hydro1K dataset. Panel B) Zoom detail of raster in Panel A. The white lines are the streamlines for which the distances were calculated.

2.3.6 Vector Species Distribution

See **Chapter 1 Section 1.5** *The Simuliidae vector hosts of onchocerciasis* for an introduction to the blackfly vectors of onchocerciasis and **Appendix A Table A-1** for species distributions. Sibling species (cytospecies) of the *S. damnosum* complex in forest and savannah zones have differential capacities to transmit infective L3 larvae of *O. volvulus* (*154, 155*). Factors responsible for these differences include:

- Duration of the gonotrophic cycle
- Degree of anthropophagy (proportion of bites taken on human hosts)
- Adult female survival rate and flight range
- Parasite-induced vector mortality
- Density-dependent uptake of microfilariae (particularly in savannah members of *S. damnosum* s.l.) and development to L3 larvae within vectors

The vector species most likely to be responsible for transmission in a segment of the study area was derived from a *S. damnosum* species cytotype distribution map (**Figure 2-5**) prepared by Meredith and Davies (1981, unpublished) and compiled from published data and internal OCP reports. This

map was digitised, geo-referenced and overlaid onto the study area. Points of identified species cytotypes from this map were then digitised and labelled using this overlay.

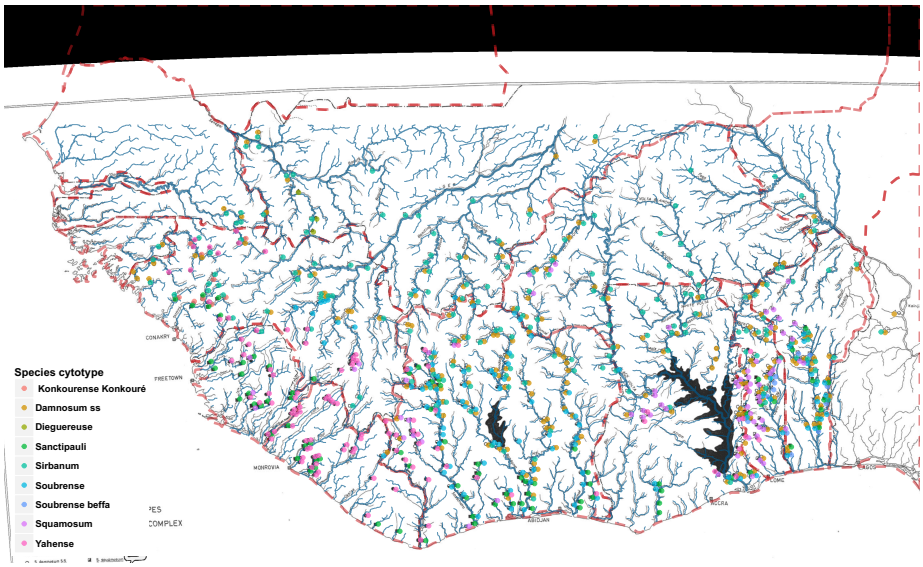


Figure 2-5: Digitised species identification map.
A hard-copy paper map was scanned in and then geo-referenced using the ArcGIS geo-referencing toolbar (ArcGIS 2001) (*156*). The streamlines from the Hydro1K dataset were overlaid over the streamlines depicted in the paper map, after geo-referencing. There is good alignment between the digitally transformed and geo-referenced paper map and the streamlines derived from satellite data.

Thiessen or Voronoi polygons (*157*) were constructed around each species identification point from the map, such that the polygonal boundary surrounding a given point encompasses the area closer to that point than any other data point in the dataset (**Figure 2-6A**). The main vector species responsible for transmission within each polygon was assumed to be the same as the species identified at each point. In addition, vector species may be classified as to whether they are forest-type or savannah-type species. This distinction is shown in **Figure 2-6B** showing good concordance with the known distributions of forest and savannah vector species, given the distribution of forest zones in the OCP area.

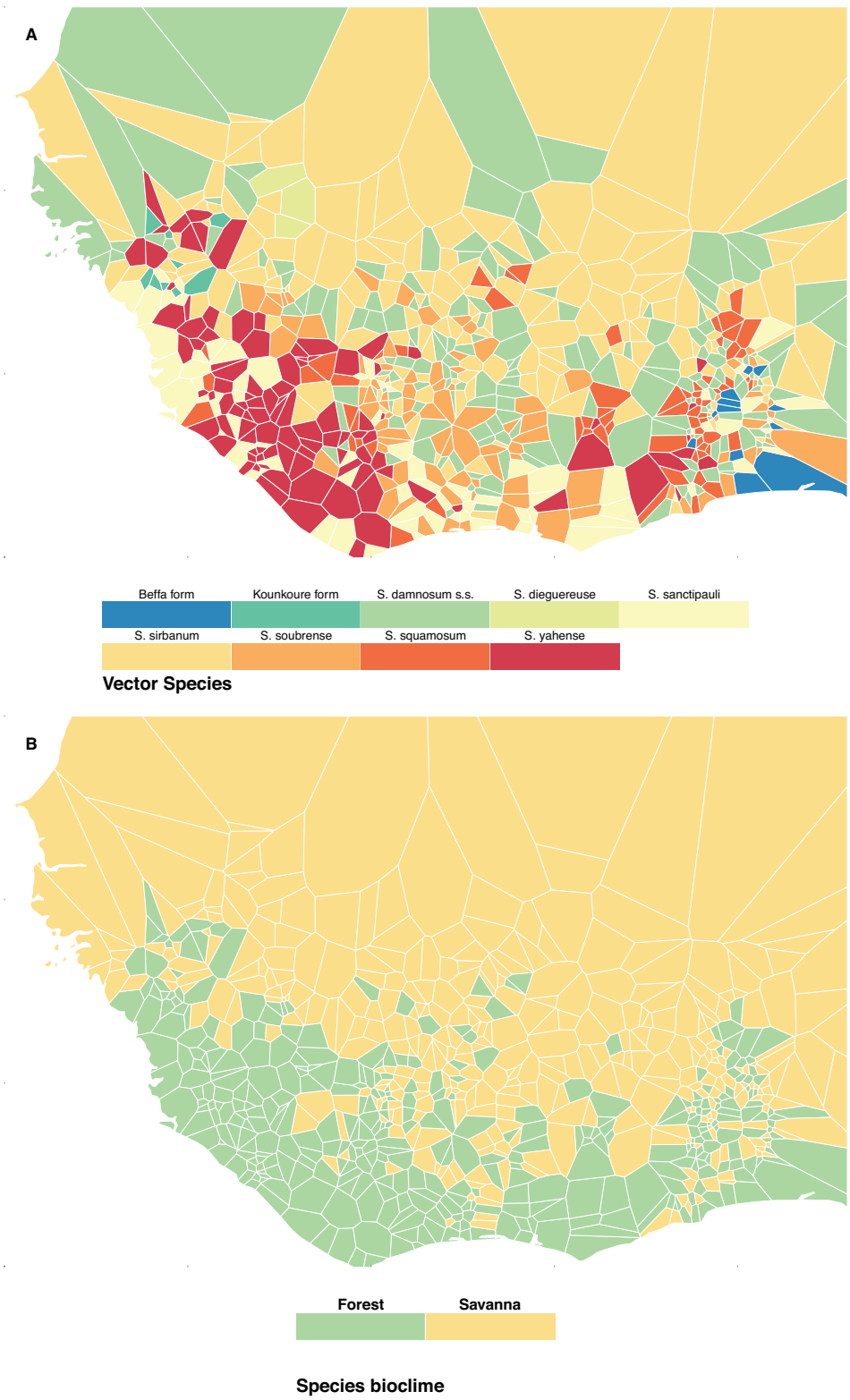


Figure 2-6: Voronoi tessellations of Simulium species.
Points were digitised from Meredith and Davies (1983, unpublished) species distribution map (**Figure 2.5**). Panel A) Polygons are coloured by species name. Beffa form and Konkoure form refer to different cytoforms of *S. soubrense*. Panel B) Polygons are coloured by species habitat bioclime (forest or savannah).

2.3.7 Initial exploratory analysis of spatial autocorrelation

An initial exploratory analysis was carried out to assess if there was any spatial autocorrelation exhibited by the data. Spatial autocorrelation violates the assumption of independent identically distributed outcomes, assumed by standard statistical approaches to model fitting and analysis. However, by adjusting for and characterising spatial autocorrelation it can be taken advantage of to estimate values at un-sampled locations.

2.3.8 Plots of prevalence against spatial location and LOESS curves

A plot of the pre-control data locations, coloured by endemicity class was inspected for clustering of villages of the same endemicity class (**Figure 2-3**). Scatterplots of microfilarial prevalence against each of the spatial coordinates (latitude and longitude) were made, and a Local Regression (LOESS) smoother (*158*) fitted to the plots to indicate trends in the response variable (*mf* prevalence) when moving west – east (**Figure 2-7A**) or south – north (**Figure 2-7B**) across the study region. When the outcome of a response variable at a given location does not depend on the separation distance to other measured locations the LOESS curve should be flat.

2.3.9 Mantel correlograms

The Mantel test (*159*) compares two independently obtained distance matrices, x and y. In the application to geostatistics, matrix x may be a matrix of ecological distances for some response variable obtained between all pairs of n survey points. Matrix y is a matrix of geographic distances between all pairs of n survey points. The Mantel statistic is calculated by computing the correlation between x and y according to the normalized Mantel statistic,

$$r = \frac{1}{(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{(x_{ij} - \bar{x})}{s_x} \frac{(y_{ij} - \bar{y})}{s_y}, \text{ for } i \neq j \tag{2.1}$$

Where i and j are row and column indices of the similarity matrices, n is the number of sampling locations for which geographic and ecological distances have been calculated and values of s are the standard deviations of the distances in matrices x and y, respectively (*160*). The statistic obtained is compared to the distribution of Mantel statistics obtained through random permutations of the data. If there is no spatial autocorrelation the expectation is for the Mantel statistic obtained from the true data to lie near the centre of the distribution of Mantel statistics obtained from random permutations. In the presence of spatial autocorrelation, at a given significance level, the obtained statistic will be more extreme. Probability values are adjusted for multiple testing according to the method of Holm (*161*).

The correlogram is obtained by computing a Mantel statistic for response variables and sampling stations in increasing distance classes. The ecological distance matrix may be derived directly from the measured response at the n sampling locations, or it may be derived from the residuals of a model fitted to the data. Mantel correlograms were computed from microfilarial prevalence at sample locations, and for residuals from fitting first and second order polynomial trend surfaces. They were also computed for univariable models for distance to nearest river, vector species type and hydrologic rank, as well as models with interaction terms. For details of models used to calculate residuals see legend of **Figure 2-8A-F**.

2.3.10 Spatial weight matrices

Fundamental to all autocorrelation statistics is the weights matrix, used to define the spatial relationships of the regions or nodes of a connected graph or neighbourhood, so that (usually) those that are close in space are given greater weight in the calculation of an outcome than those that are distant from each other (*162*). To test for spatial autocorrelation one must first generate a matrix of spatial weights, which represent spatial interaction between geographic regions, or in this case, points

in Euclidean space. There are several ways of generating spatial weights and the choice of spatial weight matrix will have influence on the estimate of spatial dependence in the data (*163*). For this reason, the dataset was tested using several different constructions of spatial weights matrices (**Figure 2-9**). The spatial weight matrices that were used were:

- Delaunay Triangulation
- Gabriel Graph (*164*)
- Relative Neighbourhood (*165*)
- Minimum Spanning Tree
- J-Nearest Neighbours (with J=4 and J=10)

The graph types are listed in order of connectedness. The Delaunay triangulation is the most connected graph. The Gabriel graph is a sub-graph of the Delaunay triangulation (all edges in the Gabriel graph are contained within the Delaunay triangulation), the relative neighbourhood graph is a subgraph of the Gabriel graph etc.

2.3.11 Global indicators of spatial autocorrelation: Moran's *I*

Moran's *I* is one of several commonly used methods for formally assessing spatial autocorrelation (*166*) (including Geary's c, and Tango's tests), with positive spatial autocorrelation indicating the existence of clustering of areas with similar attribute values, and negative spatial autocorrelation indicating that neighbouring areas tend to have dissimilar attribute values (inhibition). Moran's I is a measure of global spatial autocorrelation (i.e. a measure of the overall clustering of the data across the whole study space), that is less sensitive to local spatial autocorrelation. Moran's *I* is defined as,

$$I = \frac{n}{W} \cdot \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \overline{x})(x_j - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \qquad (2.2)$$

Where W is the sum of all weights in the spatial weights matrix, n is the number of spatial features, i and j are indices, x is the outcome variable (mf prevalence) at a given spatial index and $\bar{x}$ is the mean of the outcome variable at the n sampling locations. The expectation of Moran's I under the null hypothesis of no spatial autocorrelation is,

$$E(I) = \frac{-1}{n-1} \tag{2.3}$$

As n becomes large, $E(I)$ tends to 0. Values greater than the expected value indicate positive spatial correlation, whilst values less than the expected value indicate negative spatial correlation. Statistics for Moran's I were calculated from residuals obtained by fitting the same models as listed in **Section 2.3.9** *Mantel Correlograms* and are given in **Table 2-2**.

2.3.12 Local indicators of spatial association: Local Moran

Local Indicators of Spatial Association (LISA), proposed by Anselin (1995) (*167*) are designed to indicate the degree of significant spatial clustering around individual observations and can illustrate hot spots of spatial clustering. The sum of LISAs should be proportional to a global indicator of association of the same type. The Local Moran statistic was computed for each location. For an observation x at location i the local Moran statistic I_i is defined as,

$$I_i = (x_i - \bar{x}_{J_i}) \sum_j w_{ij} (x_j - \bar{x}_{J_i}) \text{ , with } j \in J_i \tag{2.4}$$

where $\bar{x}_{J_i}$ is the mean of points in neighbourhood J (defined by the spatial weight matrix) of location i, and w_{ij} is as above, the weight between locations i and j. The summation of weights between i and j occurs over all locations j which are defined as neighbours of i in the neighbour J_i. (*167*).

The results are displayed as a series of scatterplots in **Figure 2-10 A-H**. The x-axis values are the residual deviation for each of the data points, with Panels A-H referring to residuals obtained by

fitting the same models as used in **Section 2.3.9** *Mantel Correlograms*. The *y*-axis values are the spatially-lagged residuals from this model, i.e. for each point *i* on the plot, the average value for residuals of neighbouring points within the neighbourhood J_i.

A regression line was fitted to these points, the slope of which is proportional and roughly equivalent to the global Moran's statistic, *I*. Each plot is divided into four quadrants. The values in the upper-right quadrant occur when high values cluster with neighbouring high values, points in the lower left occur when low values cluster with neighbouring low values (positive spatial autocorrelation), whilst the upper-left and bottom-right quadrants occur when low values cluster with neighbouring high values and high values cluster with neighbouring low values, respectively (negative spatial autocorrelation).

In the absence of any spatial correlation the random point pattern would give rise to a flat regression line, with a slope approaching 0, the limit of the expectation of Global Moran's *I* in the absence of spatial autocorrelation. Since the statistic for each observation much be compared against the statistic obtained from random permutations of the data (because the data are assumed to be correlated the number of degrees of freedom is unknown) the obtained p-values are corrected for multiple testing by the false discovery rate method (*168*).

As *P*-values are obtained for each data point, it is possible to visualise where clusters of significant local spatial autocorrelation are occurring by plotting the data points on a map, stratified by their exceedance of a given significance level. **Figure 2-11** displays locations exhibiting significant local clustering with a $P < 0.05$.

2.3.13 Polynomial trend surfaces

A simple polynomial trend surface estimates the value of a variable at a given location using only the coordinates of that location (*169*). It is algebraically equivalent to a simple linear regression on the

coordinates of the data. This smooth function is known as the trend surface and may be generalised to a polynomial of nth degree. The simplest form of surface is the first order polynomial, defined as,

$$Z(x,y) = \beta_0 + \beta_1 x + \beta_2 y + e(x,y) \qquad (2.5)$$

where the value of some response variable Z is a linear function of its x and y coordinates (with global intercept) and $e(x,y)$ is the unobserved random error at that location. The β regression coefficients are the same across the grid so that the value of Z depends only on the coordinates. The trend surface at prediction points is estimated by linear regression; therefore, all input data points contribute equally in this design. Higher order surfaces may also be estimated, e.g. the second order polynomial trend surface is defined as,

$$Z(x,y) = \beta_0 + \beta_1 x + \beta_2 y + \beta_3 x^2 + \beta_4 y^2 + \beta_5 xy + e(x,y) \qquad (2.6)$$

The general form of polynomial trend surfaces of order p is,

$$Z(x,y) = \sum_{r+s \leq p} \beta_{r,s} x^r \; y^s \qquad (2.7)$$

Whilst polynomial trend surfaces are easy to fit they do suffer from several major limitations, namely, the tendency to drift towards extreme values at borders of study regions and in areas where there are no data points.

2.3.14 Spline surface regressions

A spline is a continuous function, formed of a series of connected linear segments. The points at which the spline segments join are called knots, and at these points the spline function may change direction (slope). In this way, complex non-linear patterns in the data may be approximated by a set of linear polynomial functions. In this study, both smoothing splines and regression splines are used to approximate surfaces of microfilarial prevalence. Regression splines are subtly different from

smoothing splines. Smoothing splines use all the input data points themselves as knots, and they also include a smoothing parameter λ, which controls the trade-off between fidelity (better fit to data) and smoothness (less curvature in the spline) (*170*). The objective when fitting a smoothing spline to a one-dimensional vector of n observations is to minimise the following expression,

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - f(x_i))^2 + \lambda\int_{x_1}^{x_n}(f''(x))^2\,dx \qquad (2.8)$$

Where y_i is the observed value at location x_i, $f(x_i)$ is the spline estimate of y_i at location x_i, λ is the smoothing parameter and $f''(x)$ is the second derivative of the spline curve, integrated over all x to give the average curvature of the spline.

Positive or negative values for $f''(x)$ indicate curvature, whilst 0 indicates linearity. As $\lambda \longrightarrow \infty$ any curvature in the spline becomes infinitely penalised and only linear functions are allowed, whilst as $\lambda \longrightarrow 0$ curvature becomes penalty free and the spline curve passes exactly through each of the data points. This is known as the interpolation spline. To reduce the computational complexity of solving the minimisation function over all points, a spline regression is employed whereby the spline-smoothing minimisation problem is solved for a small sub-dataset, and a model from this is constructed for the estimation of points in the full dataset (*171*). The smoothing function $f(x_i)$, may be any polynomial function and may be extended into 2 or more dimensions. For this exercise multilevel *B*-splines were used.

Multilevel *B*-splines fit a weighted sum of spline surfaces to a set of data points. Splines are fitted at each data points, considering a range of scales of influence for neighbouring data points (the control points). At coarse scales, each control point has an influence on a greater number of data points, producing smoother surfaces at the expense of surface accuracy. At finer scales, each control point affects a shrinking neighbourhood of nearby data points, and each data point will therefore be approximated by its nearest neighbours.

The fitting is carried out in stages, starting at the coarsest approximation. The next iteration a more localised neighbourhood of control points is used to approximate a surface fitted to the residuals from the first iteration. This process continues, up to the most local neighbourhood of the input data points and the final approximation function is given as the sum of the functions through each iteration of the multilevel fitting process (*172*).

2.3.15 Inverse distance weighted interpolation

Inverse Distance Weighting (IDW) interpolation estimates values at prediction locations using weighted averages of neighbouring sample points, with the weights computed as a function of their distance from the point being estimated (*173*),

$$Z(x_i) = \frac{\sum_{j=1}^{n} w(x_j) Z(x_j)}{\sum_{j=1}^{n} w(x_j)} \tag{2.9}$$

Where the spatial weights are an inverse power function of Euclidean distance between location x_i and the surrounding locations,

$$w(x_i) = \left\| x_j - x_i \right\|^{-p} \tag{2.10}$$

with $\|\cdot\|$ denoting the Euclidean distance between points, and p the power determining the rate of decay of influence of nearby points. Higher values of p place greater weight on nearer points, whilst small powers place more weight on more distant points. The neighbour of points included in the function may be limited in several ways, e.g. through a maximum search radius, or by via a minimum or maximum number of neighbouring points to use.

IDW has some severe limitations. Due to the inverse power used to define weights, as p tends towards 0 or infinity the weights will tend towards 1 or 0 respectively, thus guaranteeing that interpolated

values will always be within the range of input values. IDW assumes a globally stationary trend across the entire study region and assumes a smooth gradient between known sampling points.

2.3.16 Surface estimation using kriging

A further method of estimating surfaces of microfilarial prevalence is to use geo-statistical methods, which explicitly model the covariance between observed spatial locations as a function of their geographic distance. The data are treated as a partial realisation of a random process (equivalently a stochastic process or random field) (*174*),

$$S(x) : x \in A \tag{2.11}$$

Following the notation of Diggle (*130*), *S(x)* is a real-valued stochastic process, observed at location *x*, in area *A*, where *A* is a fixed subset of $\mathbb{R}^a$ in positive *a*-dimensional space. The spatial index *x* varies continuously throughout the region *A* (*174*). A geostatistical model also incorporates a multivariate distribution for the random response variable *Y*, conditional on *S(·)*, the signal. The vector of Y_i responses is assumed to be conditionally independent given *S(·)*, that is, $E\left[Y_i \mid S(\cdot)\right] = S(x_i)$ (*130*).

2.3.17 Variogram Modelling

An exploratory tool and the fundamental quantity underpinning geostatistics is the semivariogram (*174-176*). Under the assumption of spatial dependence, there is a relationship between the variance of a pair of random variables (in this study microfilarial prevalence) and their separation distance, *d*. If we assume that the random process *S(x)* is intrinsically stationary, then the expectation of *S(x)* is the same everywhere,

$$E\left(S(x+d) - S(x)\right) = \mu \tag{2.12}$$

And the variance between locations separated by distance *d* is given by,

$$\mathrm{var}\big(S(x+d)-S(x)\big)=2\gamma(d) \qquad (2.13)$$

The quantity $2\gamma(d)$, which describes the degree of spatial dependence across *S(x)* is known as the variogram, the classical estimator of which was proposed by Matheron (1962) (*176*) as,

$$2\gamma(d)\equiv\frac{1}{|N(d)|}\sum_{|N(d)|}\Big(S(x_i)-S(x_j)\Big)^2 \qquad (2.14)$$

Where $N(d)=\{(x_i,x_j):x_i-x_j=d\}$ and $|\cdot|$ denotes the number of separation distances between all locations separated by *d*. More usually the quantity $\gamma(d)$ is estimated, and is half the average squared difference between points separated by distance *d* and is therefore called the semivariogram.

An extension to this framework is to allow a spatially varying mean, which may be achieved via the inclusion of explanatory covariates, e.g. replacing the constant mean with a linear model for the response variable being modelled (*130*). The empirical semivariogram can then be computed from the model residuals, which are known as the trend. In most sampling designs, observations will not be made at regularly spaced points. Therefore, to accommodate irregularly spaced observation the variogram is estimated over a series of lags. The lags are simply interval classes of distances to which point-pairs are assigned, e.g. the semivariance for all point pairs separated by distances between 0 and 100 metres. The theoretical variogram is a model fitted to the empirical variogram and is given by a covariance function, which computes the covariance between variables located by separation distance *d* and is the product of a variance parameter, σ^2 and a correlation function $\rho(d)$ (*130*) as follows,

$$C(d)=\sigma^2\rho(d) \qquad (2.15)$$

The basic parameter of the correlation function, $\rho(d)$ is ϕ which is known as the range parameter. It determines the rate of decay of spatial dependence with respect to distance, *d*. Common correlation

functions often used in geostatistical applications include the exponential, Gaussian, spherical, and Matern functions (among others not discussed here), whose expressions are listed here:

Exponential: $$\rho(d) = \exp(-\frac{d}{\phi}) \tag{2.16}$$

Gaussian: $$\rho(d) = \exp\left[-(\frac{d}{\phi})^2\right] \tag{2.17}$$

Spherical: $$\rho(d) = \begin{cases} 1 - 1.5\frac{d}{\phi} + 0.5(\frac{d}{\phi})^3 & : if\ d < \phi \\ 0 & : if\ d > \phi \end{cases} \tag{2.18}$$

Matern: $$\rho(d) = \frac{1}{2^{\kappa-1}\Gamma(\kappa)}(\frac{d}{\phi})^{\kappa}\mathrm{K}_{\kappa}(\frac{d}{\phi}) \tag{2.19}$$

Where Γ is the gamma function, which gives the factorial of an integer *n* shifted by one such that $\Gamma(n) = (n-1)!$. The function K_{κ} is the modified Bessel function of the second kind, with order κ which is known as the smoothness parameter.

The theoretical variograms given by the covariance functions above are described by the following parameters:

- The nugget (τ^2): most variograms do not pass through the origin. This discontinuity at zero separation distance is known as the nugget variance and may represent measurement error, or a sampling design that is too coarse to capture spatial sources of variation at very small distances.
- The sill (σ^2): the maximum semivariance, or height of the variogram. The parameter σ^2 is often referred to as the partial sill whilst $\tau^2 + \sigma^2$ is the true sill, which is the notation used here.

- The range (ϕ): the distance at which the sill variance is obtained. In some models, e.g. the exponential model, the variogram asymptotically approaches the sill variance, so the range is given as the effective or practical range, and is usually defined as the distance at which the variogram reaches 95% of the sill variance.

Given the exponential correlation function above, an exponential variogram model, including a nugget effect, is specified by the equation,

$$\sigma^2\left[1-\exp(-\frac{d}{\phi})\right]+\tau^2 \qquad (2.20)$$

The correlation functions listed above were each tested for goodness of fit to the data. After consideration of parametric assumptions of the models under review, the best fitting correlation function was selected and used in further kriging methods. Models were fitted using maximum likelihood estimators (MLE) [**Table 2.4**] and, in addition, via minimised weighted least squares (MWLS) [**Table 2.5**]. The loss function is given by $MWLS(\theta)=\sum_k n_k\left[(\hat{\gamma}_k)-\gamma_k(\theta)\right]^2$ where θ denotes the vector of parameters of the correlation function and for each *k*th lag class; n_k denotes the number of pairs in that class, and variable $(\hat{\gamma}_k)$ denotes the value of the empirical variogram, whilst $\gamma_k(\theta)$ is the value of the theoretical variogram (*177*). Initial parameter values for the numerical optimisation algorithm were found by testing a matrix of initial starting values for the range parameter (ϕ), the sill ($\tau^2+\sigma^2$) and the nugget variance parameter (τ^2).

The MLE and the MWLS parameters were calculated for each combination of initial values, and the set that, either maximized the likelihood (MLE) or minimized the loss function (MWLS), was used as a starting point for numerical optimization of model parameters. MLE model fitting was carried out using the likfit() function from the geoR package (version 1.7.4) (*130, 177*) using R (version 3.02) (*178*). Because MLE maximises likelihood given the data, all point pairs were used in the estimation

of model parameters (known as a variogram cloud), and therefore an empirical semivariogram (which is just an aggregation of the data based on distance classes or lags) is not needed. MWLS model fitting was carried out using the variofit() function from the geoRglm package (version 0.9.4) (*179*).

2.3.18 Simple kriging

Simple kriging (SK) estimates the value of a known response Y_i at location i, as a weighted sum of the response at surrounding data points, for a known mean. More formally, SK is the minimum mean square error prediction under the stationary Gaussian model,

$$\hat{S}(x) = \mu + \sum_{i=1}^{n} w_i(x)(y_i - \mu) \tag{2.21}$$

Continuing in the notation of Diggle (*130*), the minimum mean square error predictor, $\hat{S}(x)$ of the signal $S(x)$ at arbitrary location x is a function of the observed responses $y = (y_i, \ldots, y_n)$ which minimises $E\left[\left\{\hat{S}(x) - S(x)\right\}^2\right]$. The $w_i(x)$ are the kriging weights at each location which are a function of the covariance parameters ϕ, σ^2 and τ^2. A map of mean microfilarial prevalence and the kriging variances using SK are given in **Figure 2-17**.

2.3.19 Ordinary Kriging

Whereas SK assumes a known stationary mean, with $E\left[S(x)\right] = \mu$, ordinary kriging (OK) estimates the mean $\mu(x_i)$ at an arbitrary location x_i, given the observed values in the neighbourhood of x_i (the neighbourhood could extend to the entire set of points over which Y_i is observed). The OK estimator is given by,

$$\hat{S}(x) = \sum_{i=1}^{n} w_i(x)\mu(x_i) \tag{2.22}$$

The weights are chosen such that $E\left[\hat{S}(x)\right]$ equals $\mu(x)$. Since $\mu(x)$ and $\frac{1}{n}\sum_{i=1}^{n}\mu(x_i)$ are assumed to be equal, the weights are computed such that they sum to 1, thus giving,

$$\sum_{i=1}^{n} w_i(x) = 1 \tag{2.23}$$

The advantages of OK over SK are that OK is not as affected by non-stationarity of the response variable. In SK, at locations distant to any data-points, the estimator tends towards the global average. In OK, the estimator tends towards the local average of surrounding values, an intrinsically more desirable property in spatially structured models. **Figure 2-18** display maps of microfilarial prevalence estimated by OK and the associated kriging variance respectively.

2.3.20 Universal/external trend kriging

Lastly, external trend kriging (KTE) or universal kriging (UK) are a further extension to the kriging estimator that fit a spatially varying trend surface, which may be a polynomial on the x-y coordinates of the prediction locations, or incorporating a model for the response variable at unknown locations, based on a linear combination of some explanatory covariates. **Figure 2-19** display maps of microfilarial prevalence estimated by KTE and the associated kriging variance respectively.

For a full explanation of KTE and generalised linear geostatistical modelling, see *Chapter 3*. In this chapter, exploratory surfaces of microfilarial prevalence incorporating an external trend for known risk factors for onchocerciasis where appropriate, were computed to test the utility of this modelling approach against the simpler alternatives outlined here.

2.3.21 Performance and validation of surface estimation methods

To estimate the utility of these approaches to modelling pre-control microfilarial prevalence in the OCP, several validation statistics were calculated from a *K*-folds cross validation (CV). Initially, data points were randomly partitioned into *K*=10 equal sized validation data sets, such that each input data point was used in the CV once. Refer to **Chapter 3 Section 3.12** for full details.

Additionally, to compare the ability of the methods implemented in this chapter to estimate microfilarial prevalence in areas where there are no data (e.g. to apply these models to APOC areas), a spatially clustered partitioning of the data into *K*-folds was also implemented. To partition data the nearest neighbourhood of *J* points was selected from a randomly chosen starting village. To use each point in the CV once, the size of each partition should be equal to $1/K$. Therefore, when $J < 74$ (with 737 points, *J*=74 represents 10% of the total input data) additional randomly selected data points were chosen to create a validation set representing 10% of the total data. Subsequent partitions were created in the same way, using remaining data that had not previously been included in a validation partition. In this study methods were validated using a *J*=30 and *J*=74 clusters of *J*-nearest neighbours (*J*-NN).

Figure 2-20 displays the corresponding plots for a randomly partitioned set of *n* validation points, with *n*=74. **Figure 2-21** displays the linear correlation between observed and estimated microfilarial prevalence at spatially-clustered validation points in a *J*-nearest neighbours CV, with *J*=74. **Table 2.7** gives the validation statistics from all models used in external trend kriging. The values given are the mean values across the 10 validations runs and include the linear correlation, the mean absolute prediction error and the mean prediction error for a validation using *J*=30 nearest neighbours, *J*=74 nearest neighbours and the *n*=74 randomly selected data points.

2.4 Results

2.4.1 Plots of prevalence against spatial location

Initial plots of microfilarial prevalence against each of the coordinates (**Figure 2-7**) revealed that there was a trend in prevalence in the south-north direction, with prevalence peaking at latitudes around 500-600Km north of the equator, decreasing as data were sampled further north. At the northerly limits of the OCP area the conditions are substantially more arid and desert-like, for larger parts of the year, with streams and rivers presenting as dry riverbeds or 'wadis', except for where the Niger River winds through Mali and Niger.

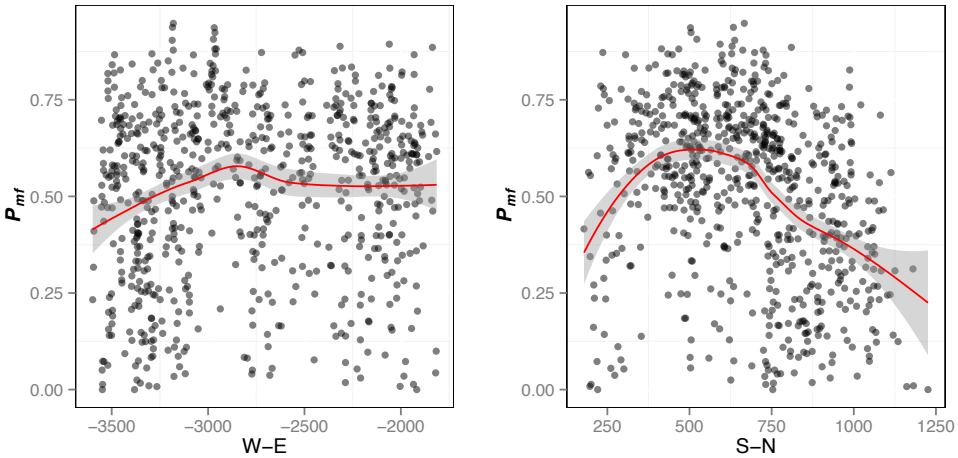


Figure 2-7: Plots of data coordinates against microfilarial prevalence.
Left panel) Microfilarial prevalence versus longitude. Right panel) Microfilarial prevalence versus latitude. The red line in each plot is a LOESS smoother with the grey shaded area around the line indicating the point-wise 95% confidence interval.

2.4.2 Mantel Correlograms

Mantel correlograms (**Figure 2-8 A-H**) show significant spatial correlation between geographical and ecological distances (pairwise differences of model residuals) under a variety of different model assumptions, at locations up to 400Km apart. There is also significant negative spatial correlation, when modelling the residuals of a 2nd order mean trend surface based on the spatial indices of the data. Significant positive spatial autocorrelation is exhibited by models incorporating environmental covariates (**Figure 2-8 D-H**) at distances greater than those exhibited by either the model that

assumes the mean is stationary (*Figure 2-8A*) or a simple function of the data coordinates (**Figures 2-8 B-C**).

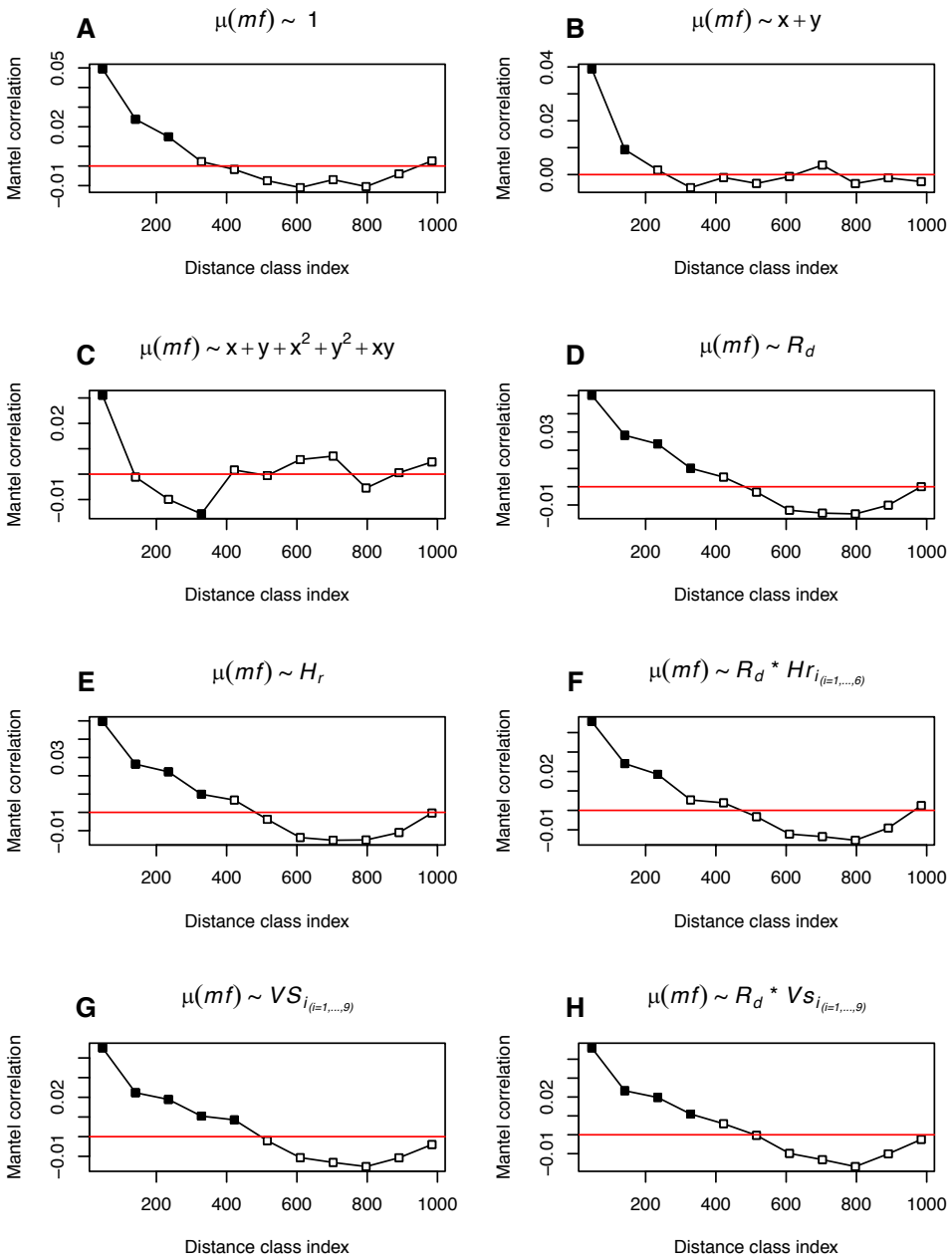


Figure 2-8: Mantel correlograms of spatial autocorrelation.
Black squares indicate significant correlation between ecological and geographical distances. The red abscissa represents the expectation of the Mantel statistic under no spatial autocorrelation. Values above the abscissa indicate positive correlation, whilst points below indicate negative correlation. The models for the residuals used in the ecological distance matrix are given above each plot, with 1 indicating an intercept only model (model A), x^p and y^p indicating a polynomial trend of order p (models B and C) fitted using a simple linear regression model. Models D through H were fitted using a negative binomial link function using a generalised linear model. Rd is a continuous covariate for distance to nearest perennial river stream, Hr is a categorical variable for hydrologic rank (Strahler order) with 6 levels, and Vs is a categorical variable for vector species with 9 levels. A '*' indicates factor crossing with the categorical variables specified, thus interaction terms, in addition to the normal additive terms for all variables in the specified model, are included. μ(mf) is the model estimated mean prevalence of microfilaria (Figure A-C) or the logit transformed estimate of microfilarial prevalence (Figures D-H).

2.4.3 Clustering With Global Moran's I

Global tests for spatial autocorrelation in the data were carried out using Moran's *I*, using a variety of spatial weighting graphs (**Figure 2-9**) to assign relative importance of near-neighbour values when assessing autocorrelation within the data. Each spatial weight graph was used to calculate Moran's *I* for each of the eight models as described in the legend of **Figure 2-8**. The results are listed in **Table 2.2**. All model and spatial weight graph combinations exhibited significant positive spatial autocorrelation (like-values clustering together), with P-values of <0.001 in each case. Multiple models and graphs were tested as it is known that the result of Moran's *I* can be sensitive to the input graph that was chosen. The model displaying the greatest positive spatial autocorrelation was that in which distance to the nearest perennial stream segment was used as a continuous predictor in a logit model (model D).

	Model	Delaunay	Gabriel	Relative neighbours	Minimum spanning tree	*J* nearest neighbours (*J*=4)	*J* nearest neighbours (*J*=10)
A	$\mu(mf) \sim 1$	0.453	0.498	0.533	0.566	0.482	0.374
B	$\mu(mf) \sim x + y$	0.395	0.45	0.489	0.524	0.429	0.314
C	$\mu(mf) \sim x + y + x^2 + y^2 + xy$	0.338	0.412	0.45	0.486	0.378	0.258
D	$\mu(mf) \sim Rd$	0.455	0.488	0.518	**<u>0.543</u>**	0.473	0.365
E	$\mu(x) \sim Hr_i : i = 1,\ldots,6$	0.454	0.487	0.516	0.542	0.473	0.366
F	$\mu(x) \sim Hr_i + (Rd \cdot Hr_i) : i = 1,\ldots,6$	0.429	0.467	0.497	0.524	0.448	0.337
G	$\mu(x) \sim V_s : s = 1,\ldots,9$	0.421	0.455	0.489	0.511	0.441	0.328
H	$\mu(x) \sim V_s + (Rd \cdot V_s) : s = 1,\ldots,9$	0.427	0.457	0.485	0.507	0.444	0.33

Table 2-2: Global Moran's I spatial autocorrelation statistics.
All models tested under all listed assumptions of neighbour relatedness resulted in significant spatial autocorrelation (all *P*-values were <= 0.001 indicating a rejection of the null hypothesis of no spatial autocorrelation). See legend of **Figure 2.9** for details of the model symbols used. The response variable for each model $\mu(x)$, is the estimated mean prevalence for models A-C and the logit transformed *mf* prevalence thereafter. The model and spatial weights graph exhibiting the greatest positive spatial autocorrelation is highlighted in bold and underlined.

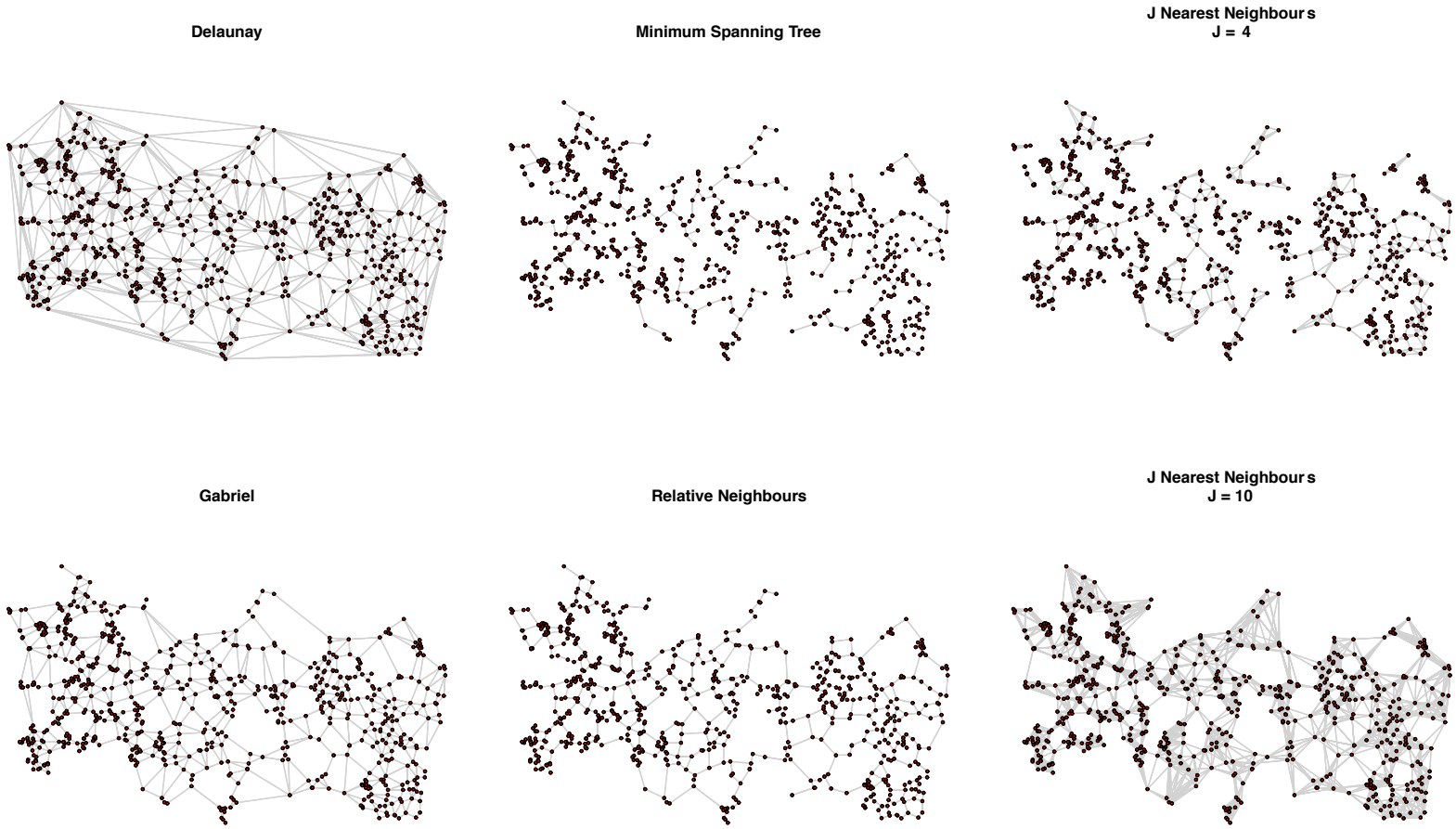


Figure 2-9: Spatial weight graph constructions.
Neighbour joining graphs for the pre-control data points used in this thesis. The red nodes represent the pre-control villages and the grey lines between them show assumed linkage. Refer to **Section 2.3.10** *Spatial Weight Matrices* for further details.

2.4.4 Local Moran's *I* and spatial location of significantly clustering data points

Local indicators of spatial association, investigated by computing a local Moran's *I* for each data point (I_i), are displayed in **Figure 2-10A-H**. The statistic I_i, was computed from model residuals. The panel labels, A-H correspond to the model used for microfilarial prevalence as given in **Table 2-2**. As a pragmatic choice, the spatial weights graph selected for the calculation of I_i was the minimum spanning tree as it exhibited the greatest average spatial autocorrelation of all the graph types across all the models ($\overline{I} = 0.525$).

In all cases, there was a trend of positive spatial autocorrelation. A linear regression of model residuals against the spatially-lagged values of residuals shows a positive trend in all cases, as would be expected given the statistics returned for the Global Moran's I. Individual data points showing significant local clustering (corrected P-value ≤ 0.05) with neighbouring data points are coloured red (high microfilarial prevalence clustering with high microfilarial prevalence ['high-high']) and blue (low microfilarial prevalence clustering with low microfilarial prevalence ['low-low']).

The spatial locations of significant points are displayed on the maps in **Figure 2-11 A-H**. Clusters of high-high points tend towards the middle latitudes of the OCP area. Clusters of low-low points tend towards the more extreme latitudes of the OCP area. As may be expected from the nature of the test, significant high-high points and low-low points tend to appear as neighbours or in clusters on the map. Maps of this nature may be useful for sampling the values of environmental covariates at these significant clusters to look for associations between those covariates and high-high or low-low clusters of villages.

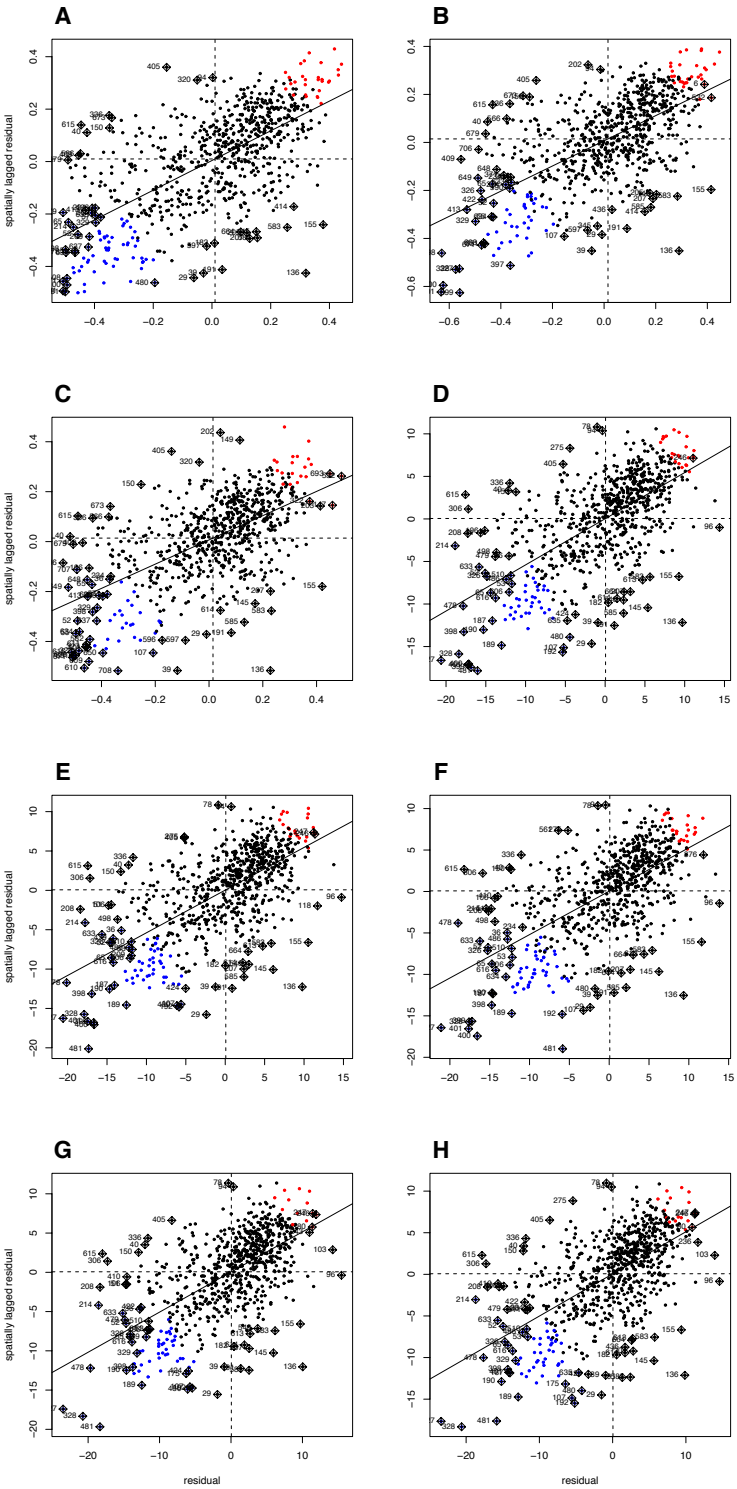


Figure 2-10: Local Moran's I scatterplots.
Values on the *x*-axis correspond to residuals of microfilarial prevalence (panels A-C) or logit transformed microfilarial prevalence (panels D-H), against their spatially lagged equivalent. Spatially lagged is simply the weighted average of neighbouring points, with weights determined by structure of the minimum spanning tree graph of inter-point relatedness. The vertical and horizontal dashed lines represent the mean residual and spatially-lagged residual values, respectively. The black solid line is a linear regression of the residuals against their spatially-lagged equivalents, the slope of which approximates the global Moran's *I*. Red points represent high predicted values of microfilarial prevalence clustering with neighbouring high predicted values of microfilarial prevalence ('high-high'). Blue points are the 'low-low' equivalent.

Figure 2-11: Locations of points exhibiting significant local clustering.
Points on these maps are coloured according to the significance of the local Moran's I_i statistic calculated for each point. Red and blue points are significant, with red denoting high estimated values clustering with neighbouring high estimated values, whilst blue are the equivalent for low estimated values. Residuals are derived from the models presented in **Table 2-2** and the legend of **Figure 2-9** *Mantel Correlograms Of Spatial Autocorrelation*.

2.4.5 Polynomial Trend Surfaces And Spline Fitting

Initial exploratory analysis has shown the data exhibits inherent spatial autocorrelation. As a first approximation, surfaces of microfilarial prevalence were interpolated using a variety of techniques, as these are relatively simple and computationally cheap to construct. For processes with a simple unobserved underlying data generating processes for the response, they may also explain the observed spatial patterns adequately.

Polynomial trend surfaces for pre-control microfilarial prevalence were calculated and are displayed in **Figure 2-12**. There are extremely simple methods and produce very smooth surfaces across the global neighbourhood of the input data points, even with high degree polynomials, which contain many parameters [the number of parameters of a surface of order p is $(p+1)(p+2)/2$]. These models were fitted by ordinary least squares and as such do not take account of the covariance structure that is present in the data. The surface they produce is very smooth, even with high degree polynomials, and an unrealistic approximation of microfilarial prevalence, with estimated values that drop below 0 (grey shaded areas on maps of **Figure 2-12**). Polynomial trend surfaces will not be considered further as a suitable approach for estimating *mf* prevalence.

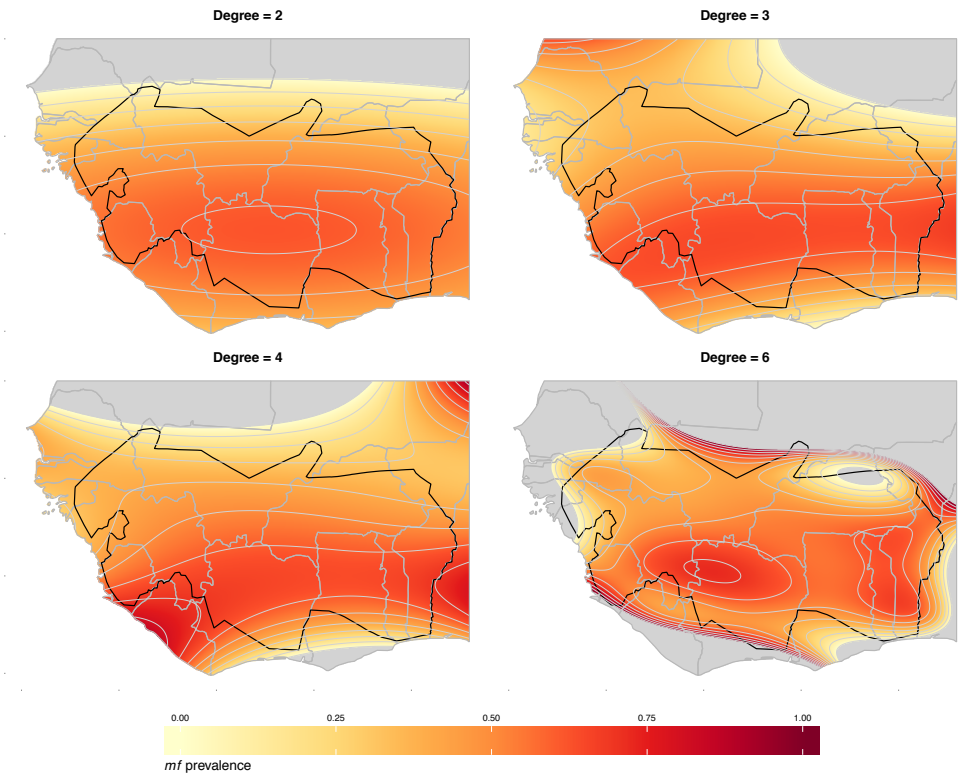


Figure 2-12: Polynomial trend surface of microfilarial prevalence.
Contour lines are displayed at 10% prevalence intervals. Grey areas are those where the surface has become unbounded by the limit, $0 \geq mf \leq 1$, where *mf* is microfilarial prevalence. The degree of the polynomial used to compute each surface is displayed above each panel.

Multilevel B-spline surfaces are displayed in **Figure 2-13**. In the absence of information from nearby data-points multilevel B-spline surfaces tend toward the global average, as can be seen in the northern stretches of the map, particularly in the area just north of the OCP borders in all maps. In areas where there are data points, multilevel B-spline surfaces are able to visualise complex patterns of variation. However, as they lacks a covariance structure and explanatory covariates, the observed pattern is not well *explained* by multilevel B-spline surfaces and therefore are not adaptable to areas lacking data. The mean prediction error (average difference between observed and expected values), the mean absolute prediction error (the magnitude of difference between observed and expected values) and the linear correlation between observed and estimated values from the validation exercises are given in **Table 2-3**. The performance of multilevel B-splines is comparable to that of more sophisticated mapping procedures when there is a good, even distribution of input data points across the x-y surface. When the density of nearby points is reduced, as in the *J*-nearest neighbours cross validation,

the validation statistics drop dramatically. The mean Pearson's correlation across all bootstrap runs in the *J*-nearest neighbours cross validation is 0.27, and included negative correlations.

Method	MPE	MAPE	Corr
N random points	-0.004	0.124	0.686
(*N=74*)	(-0.036-0.028)	(0.101-0.158)	(0.547-0.795)
K-nearest neighbours	0.023	0.177	0.270
(*K=74*)	(-0.047-0.179)	(0.131-0.234)	(-0.193-0.565)

Table 2-3: Validation statistics from multilevel B-spline surface fitting.
MPE refers to mean prediction error, MAPE refers to mean absolute prediction error and Corr refers to the Pearson correlation. The mean of the 10 bootstrap runs for either *K*-nearest neighbours or *N* random points is given. The range across the bootstrap runs is given in brackets below.

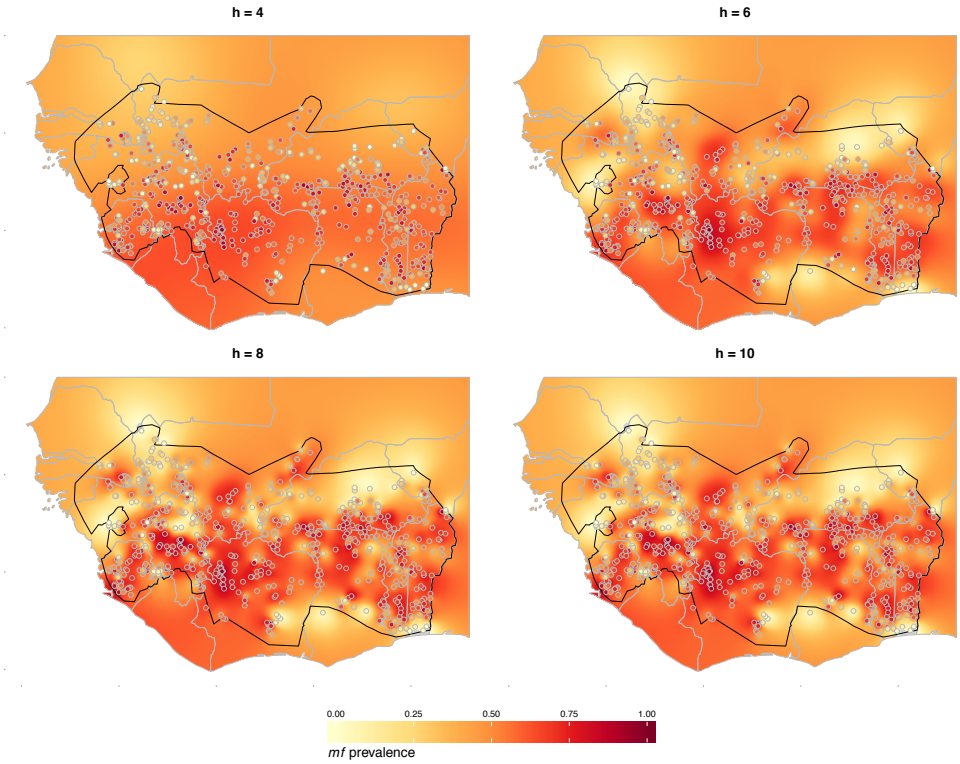


Figure 2-13: Multilevel B-spline surfaces of mf prevalence.
The *h* parameter, which controls the smoothness of the surface, is displayed above each panel. The pre-control data points are plotted as circles, coloured by endemicity and surrounded by a grey border. As the roughness of the surface increases with *h*, the data points become more difficult to see due to the focal nature of the surface, however the same points are displayed on each plot in the exact same way.

2.4.6 Microfilarial prevalence estimation with kriging techniques

Initial exploratory analysis has shown the need for a spatially structured model. Kriging methods were implemented to model and account for the spatial autocorrelation exhibited by the data. Initial variogram modelling was followed by kriging using a number of kriging methods, including external trend kriging where explanatory covariates, available at all prediction locations, where used to augment the information from the spatial covariance structure.

2.4.7 Theoretical Variogram Estimation

A number of correlation functions were used to fit theoretical variogram models to sample variograms. The simplest assumed a stationary mean across the study area (model A in **Table 2-4**). Subsequent models included polynomial trends on the coordinates (models B and C in **Table 2-4**) and models with an external trend for the linear predictor (models D-H in **Table 2-4**).

The maximum log-likelihood estimates from all correlation functions, used to fit theoretical variogram models to the empirical semivariograms under different model assumptions, are given in **Table 2-4**. Both the Matern model and the exponential model gave the same and best fit to the data, for each mean trend model.

A pragmatic choice to use the exponential correlation function was made, as it contains fewer parameters, and the exponentially decaying relationship between distance and semivariance has a plausible biological basis (*151*). However, unexpectedly, the mean trend model giving the best fit to the data was the 2nd order polynomial trend surface, the input to which is just a function of the values of the coordinates at data locations. An explanation for this could be that the other models specified do not well explain the data observed.

Alternatively (or indeed concomitantly), this could occur because the MLE optimisation process considers all pair-wise distances across the dataset. It is known that at distances approaching the

maximum lag distance (maximum pairwise separation distance between observed points) in the variogram, the spatial relationship breaks down and variogram estimates become unstable (semivariance against distance of the empirical semivariogram may fluctuate widely at high lag distances, and these points, unrelated to the spatial process, have a disproportionate effect on model fits) (*180*).

Model		Matern	Exp	Gaussian	Spherical	Circular	Cubic
A	$\mu(mf) \sim 1$	250.45	250.45	229.05	240.97	239.09	229.33
B	$\mu(mf) \sim x + y$	254.38	254.38	236.19	242.42	240.14	236.95
C	$\mu(mf) \sim x + y + x^2 + y^2 + xy$	265.51	265.51	248.45	246.58	245.22	249.22
D	$\mu(mf) \sim Rd$	256.19	256.19	236.11	243.99	242.85	236.43
E	$\mu(x) \sim Hr_i : i = 1,\ldots,6$	252.80	252.80	231.90	243.57	241.61	232.21
F	$\mu(x) \sim Hr_i + (Rd \cdot Hr_i) : i = 1,\ldots,6$	258.92	258.92	239.57	248.72	245.62	239.96
G	$\mu(x) \sim V_s : s = 1,\ldots,9$	255.49	255.49	233.68	247.50	247.10	233.68
H	$\mu(x) \sim V_s + (Rd \cdot V_s) : s = 1,\ldots,9$	266.11	266.11	246.26	258.31	257.27	246.13
Mean MLE:		**257.48**	**257.48**	**237.65**	**246.51**	**244.86**	**237.99**

Table 2-4: Maximised log-likelihood values for theoretical semivariogram models fitted to empirical semivariograms. The models used to specify the mean trend of the empirical semivariograms are shown in the first column and labelled A-H. The models presented in **Table 2-2** and the legend of **Figure 2-9** *Mantel correlograms of spatial autocorrelation*. Column headers denote the correlation functions used to compute the theoretical semivariogram models.

For this reason, the theoretical variogram models were re-fitted using MWLS, with lag distances only up to 1/3 the maximum distance between data points (~600 Km). The MWLS method assigns higher

weights to earlier distance classes, penalising those lags with few observations (*180*). The values for each theoretical variogram model correlation function are presented in **Table 2.5**.

Model		Matern	Exp	Gaussian	Spherical	Circular	Cubic
A	$\mu(mf) \sim 1$	0.181	0.163	0.392	0.367	0.348	0.473
B	$\mu(mf) \sim x + y$	0.118	0.139	0.119	0.128	0.123	0.122
C	$\mu(mf) \sim x + y + x^2 + y^2 + xy$	0.472	0.482	0.446	0.443	0.443	0.444
D	$\mu(mf) \sim Rd$	0.172	0.147	0.326	0.307	0.337	0.332
E	$\mu(x) \sim Hr_i : i = 1, \ldots, 6$	0.109	0.100	0.188	0.186	0.202	0.206
F	$\mu(x) \sim Hr_i + (Rd \cdot Hr_i) : i = 1, \ldots, 6$	0.111	0.102	0.190	0.206	0.203	0.282
G	$\mu(x) \sim V_s : s = 1, \ldots, 9$	0.105	0.102	0.166	0.161	0.173	0.178
H	$\mu(x) \sim V_s + (Rd \cdot V_s) : s = 1, \ldots, 9$	0.089	**0.079**	0.155	0.151	0.163	0.228
Mean MWLS:		**0.170**	**0.164**	**0.248**	**0.243**	**0.249**	**0.283**

Table 2-5: Minimised value of weighted least-squares for theoretical semivariograms models fitted to empirical semivariograms.
These values are calculated by fitting a parametric model and minimising a loss function. The model descriptions are presented in **Table 2-2** and the legend of **Figure 2-9** *Mantel Correlograms Of Spatial Autocorrelation*.

After re-fitting to the distance-restricted data, the empirical variogram with a mean trend specified by a 2nd order polynomial model produced the worst fit across all correlation functions, indicating that the previous MLE fit values were being achieved by fitting to data points at extreme lag distances, which is not appropriate when trying to model the spatial covariance structure. The theoretical semivariogram with the best overall fit to the empirical variogram was the semivariogram with a mean trend specified by model H.

Model H included a linear predictor with distance to nearest perennial river, simuliid vector species and an interaction between those variables. The best overall correlation function across all mean trend models was the exponential function, which was selected to compute microfilarial prevalence surfaces by kriging. Pragmatically, this is justified because the exponential correlation model includes the fewest parameters of any model, in addition to producing the smallest residual sum of squares (RSS) values. Using the Akaike Information Criterion (AIC) the discriminant expression used to choose between two models reducing RSS is,

$$AIC = n \cdot \log(\frac{RSS}{n}) + 2p \qquad (2.24)$$

Where *RSS* is the residual sum of squares, *n* are the number of data points and *p* the number of parameters (*181*). Clearly, as *n* is equal across all correlation functions used, and *p* is smallest in the exponential correlation function, choosing the exponential model based on lowest RSS values is valid in this case.

The values for the range parameter (ϕ), the sill parameter (σ^2), and the nugget variance parameter (τ^2) for the exponential correlation function are given in **Table 2-6**. The values from both the MWLS and MLE methods are listed for each of the empirical semivariograms under consideration. The range parameter estimates for all models (A-H) using MLE are all below the value estimated by MWLS. In addition, the sill variance parameter is higher using MLE than MWLS. The lower value for the ϕ parameter using MLE means that the covariance model approaches the sill variance more rapidly. The semivariance values at extreme lags, where the spatial dependence breaks down, has a large effect on the model fit using MLE, and therefore the MWLS values were used for prediction, restricting the range over which the values are estimated to be 1/3 the maximum possible lag distance.

Model		Minimised Weighted Least Squares Estimation			Maximum Likelihood Estimation		
		σ^2	ϕ	τ^2	σ^2	ϕ	τ^2
A	$\mu(mf) \sim 1$	0.031	100.000	0.019	0.043	61.085	0.009
B	$\mu(mf) \sim x + y$	0.034	50.000	0.010	0.039	45.903	0.008
C	$\mu(mf) \sim x + y + x^2 + y^2 + xy$	0.034	38.805	0.008	0.035	32.404	0.006
D	$\mu(mf) \sim Rd$	0.032	88.955	0.017	0.042	57.857	0.009
E	$\mu(x) \sim Hr_i : i = 1,\ldots,6$	0.031	80.343	0.017	0.042	61.433	0.009
F	$\mu(x) \sim Hr_i + (Rd \cdot Hr_i) : i = 1,\ldots,6$	0.031	77.776	0.016	0.042	58.111	0.009
G	$\mu(x) \sim V_s : s = 1,\ldots,9$	0.030	70.000	0.016	0.042	67.327	0.010
H	$\mu(x) \sim V_s + (Rd \cdot V_s) : s = 1,\ldots,9$	0.028	78.908	0.018	0.042	67.451	0.010

Table 2-6: Theoretical semivariogram parameter estimates from exponential correlation function.
The exponential correlation function was selected for use in the kriging surface estimation model. The parameter values from the two model fitting processes in **Table 2-4** (Maximum Likelihood) and **Table 2-5** (Weighted Least Squares) are given here. The models are presented in **Table 2-2** and the legend of **Figure 2-9** *Mantel Correlograms Of Spatial Autocorrelation*.

The fitted theoretical variogram models estimated by MLE were plotted against the empirical semivariogram cloud (**Figure 2-14**). The models estimated by MWLS were plotted against the binned empirical semivariograms (**Figure 2-15**) to assess visually how the models fitted the data.

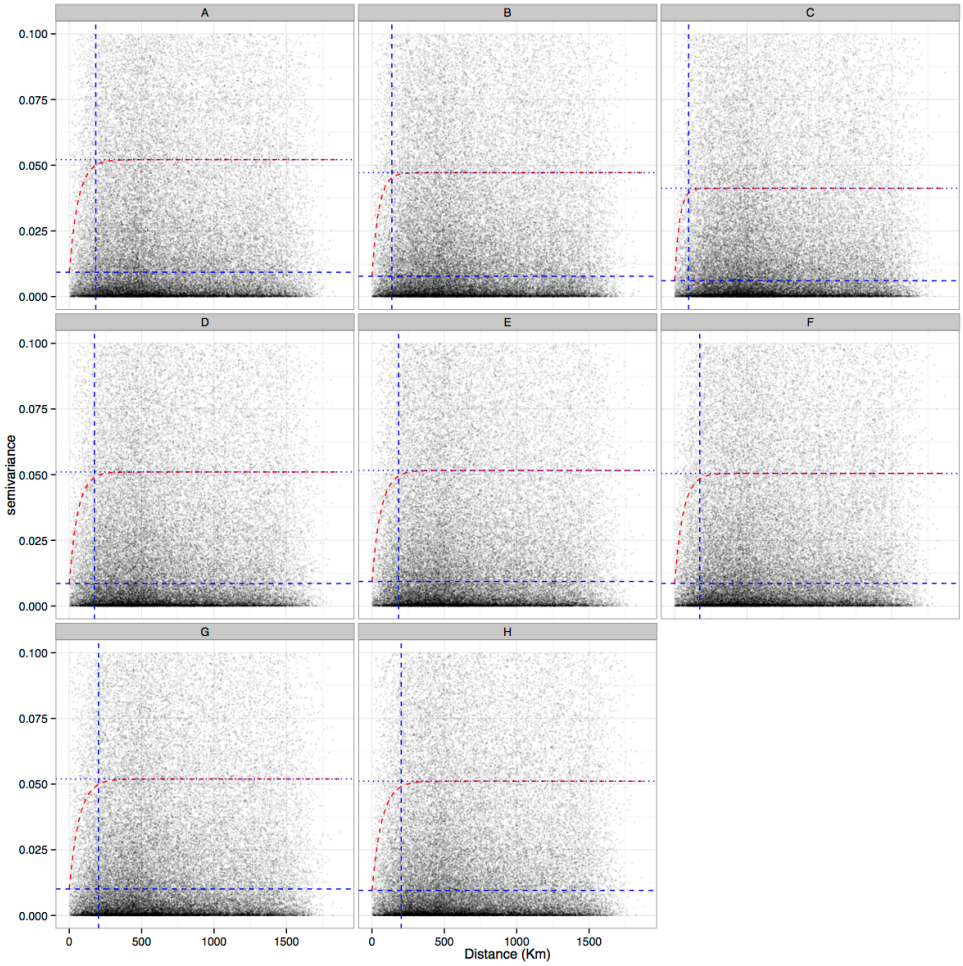


Figure 2-14: Variogram clouds with theoretical variogram model fits.
Each point represents the squared difference between pairs of points on the *y*-axis, plotted against their separation distance on the *x*-axis. The red line is the exponential theoretical variogram, with parameters estimated using maximum likelihood. The dotted blue horizontal line represents the sill (σ^2 parameter) and is the overall variance between the points. The vertical blue line represents the distance at which the theoretical variograms reach 95% of the sill value (rate controlled by the range parameter (Φ). The dashed horizontal blue line represents the nugget variance (τ^2). For display purposes $3x10^3$ points from the variogram cloud for each empirical variogram were chosen at random and plotted; however, the theoretical variogram lines were fitted using all data point pairs (n=271,216) in each model. The *y*-axis range (semivariance) was also restricted to visualise better the theoretical variogram line. A-H refer to the different models specified in **Table2-2** and the legend of **Figure 2-9**.

The MLE fit to the empirical variogram cloud treats all points equally, and thus in this setting the fits are dominated by the extreme values found at larger separation distances. The models quickly saturate to the sill variance.

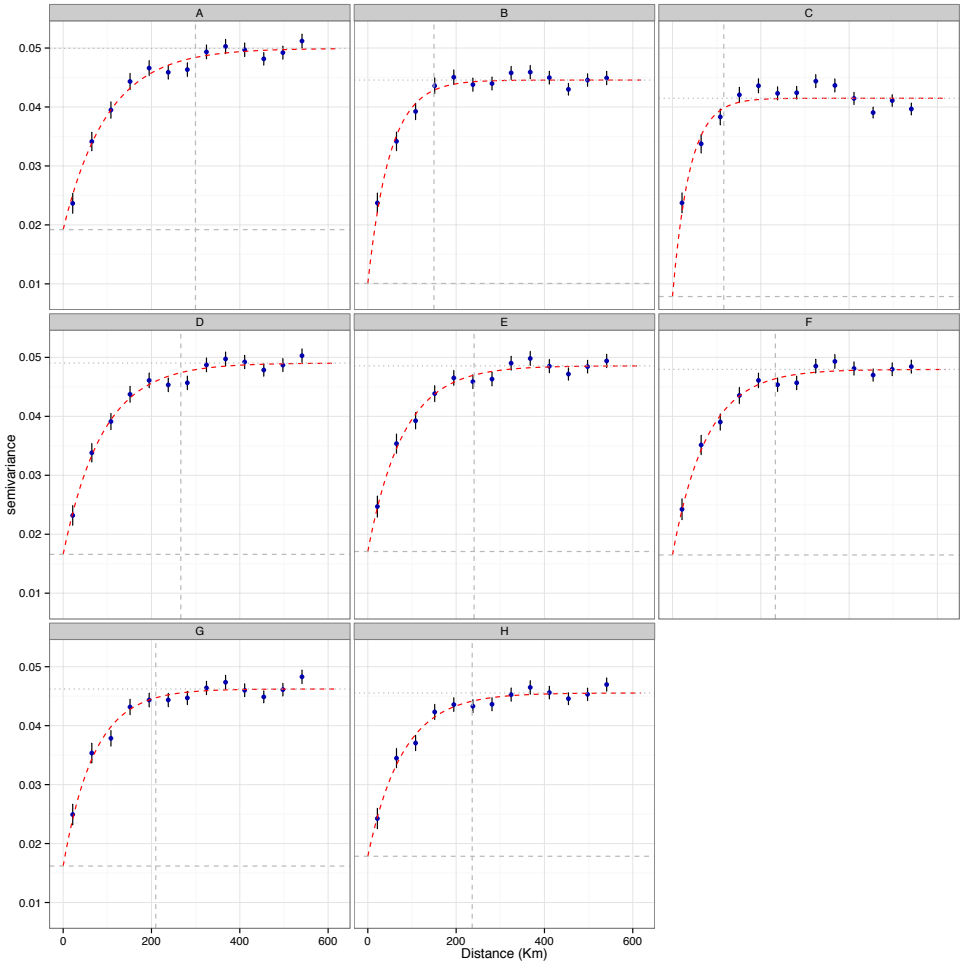


Figure 2-15: Binned sample semivariograms with fitted theoretical variogram models.
The blue points represent the mean semivariance between all point pairs falling within each distance class (the mean of each distance class or lag is the value plotted on the x-axis). The models used to calculate residuals at data points are given in the panel labels above each plot and correspond to the models labelled A-H in **Tables 2-2** and legend of **Figure 2-9**. The vertical bars show the 95% confidence interval for the semivariance of each lag class. Note the restricted *x*-axis and change of scale on the *y*-axis compared to **Figure 2-14**; only pairs of points up to a maximum separation distance of 1/3 the maximum distance between any two points were considered. At this lower range, a greater proportion of points have a low squared difference, compared to the same proportion of points at higher distances, reducing the sill variance (σ^2). See **Figure 2-16** for a sample semivariogram unrestricted by distance. The exponential theoretical variogram model fitted is the red dashed line, with parameters estimated by MWLS. The dashed vertical and horizontal grey lines represent the range (Φ) and nugget variance (τ^2) of the model respectively. The dotted horizontal grey line represents the sill variance (σ^2).

The models fit using MWLS assign a greater weight to ordinates of the empirical variogram at smaller separation distances, hence the spatial structure over shorter distances is better accounted for by disregarding spurious spatial correlation that occurs over longer distances (see **Chapter 3 Section 3.3.5** *Spatial Dependence In The Data And Model-Based Geostatistical Analysis* for semivariograms including Monte Carlo envelopes of statistically significant spatial autocorrelation). This phenomenon is illustrated in **Figure 2-16**. The use of non-weighted ordinates in the variogram to estimate model parameters, can sometimes lead to underestimation of the spatial relatedness at small lag distances.

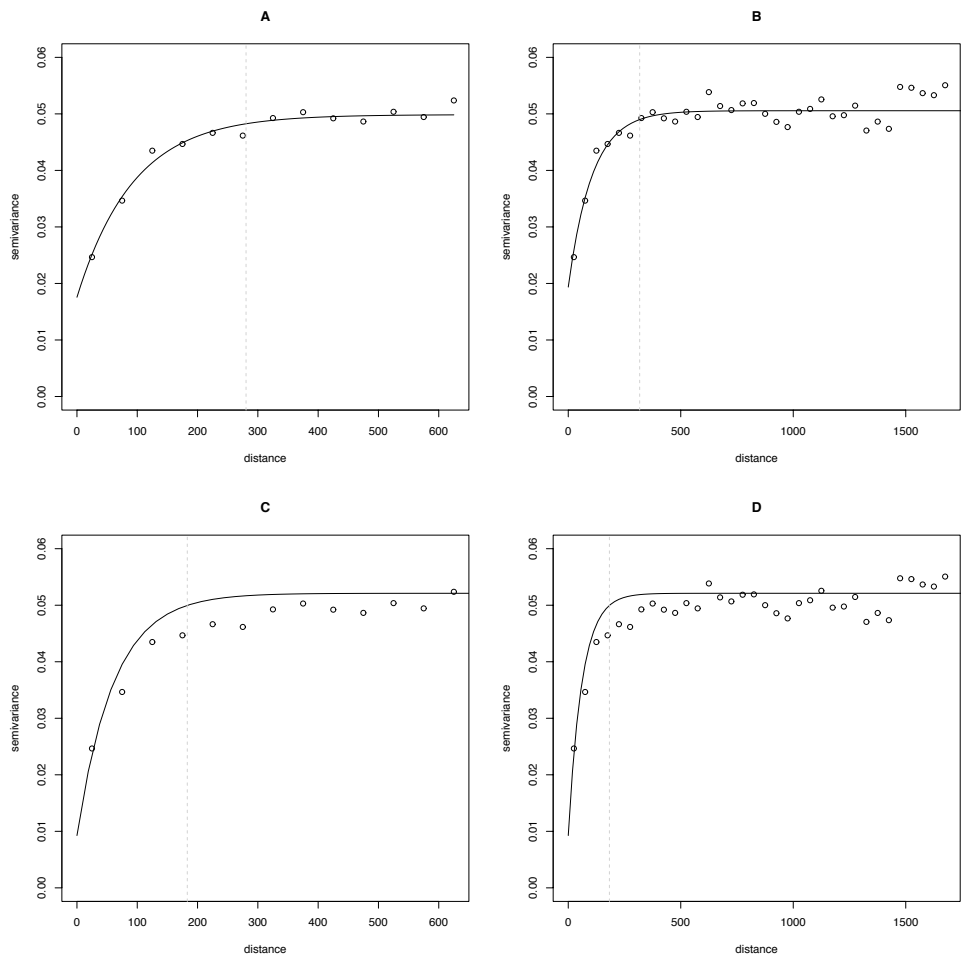


Figure 2-16: Ordinate weighting and spatial lag bias.
All panels show an exponential correlation fun used to fit a theoretical variogram model to the data. Panel A) Shows the model fit using MWLS to the sample semivariogram using data up to a maximum lag distance of 1/3 the largest distance between data points. Panel B) Shows the exponential model fitted to a sample semivariogram using all available data. At increasing distance the spatial variation becomes uncorrelated. Panel C) Displays the theoretical variogram fitted using MLE to the variogram cloud, against the sample semivariogram i.e. the line is not fitted to the aggregated data points shown in the panel. Because there is no weighting of early lags, the model quickly saturates to the overall semivariance of the data underestimating spatial correlation at small separation distances. Panel D) Shows the exponential theoretical variogram line plotted across the entire range of the sample semivariogram. Note how the line is a better fit to the later lag classes which are spatially uncorrelated, i.e. the correlation of points at these distances is random.

2.4.8 Simple kriging estimate of mf prevalence

A surface of estimated *mf* prevalence, computed using simple kriging is presented in **Figure 2-17**.

The surface is smooth and quickly approaches the global mean (which is assumed known and constant) as prediction locations move away from observed locations.

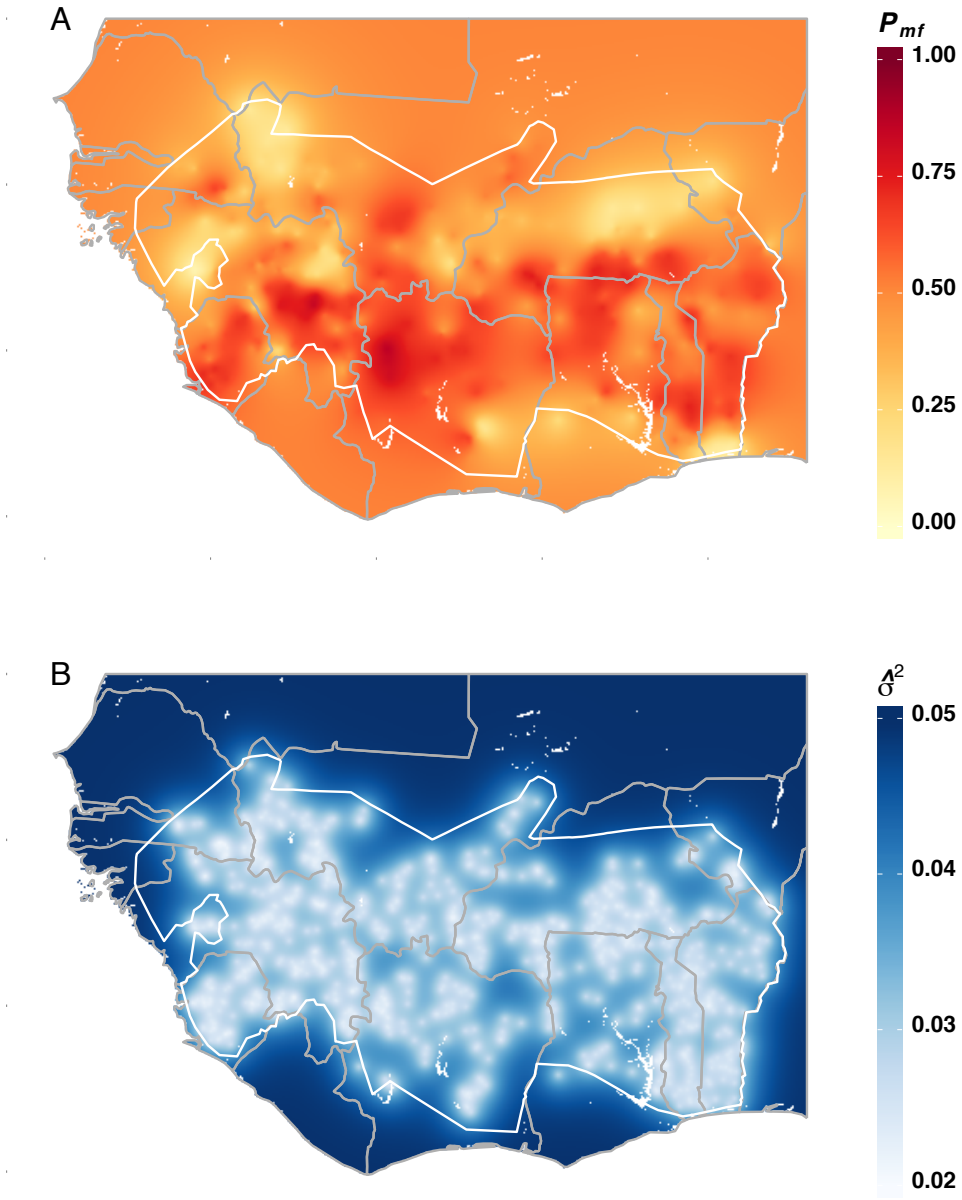


Figure 2-17: Simple kriging surface of mf prevalence.
Panel A) Mean *mf* prevalence assuming known constant mean across surface. Panel B) Estimated variance of kriging predictor.

To allow more flexibility, ordinary kriging with a locally estimated mean was implemented. The results are displayed in **Figure 2-18**. As the number of points used to estimate the local mean is increased the surface takes on a smoother shape and approaches the surface produce through simple kriging. At low numbers of nearest neighbours used to estimate the local mean, the surface becomes more focal.

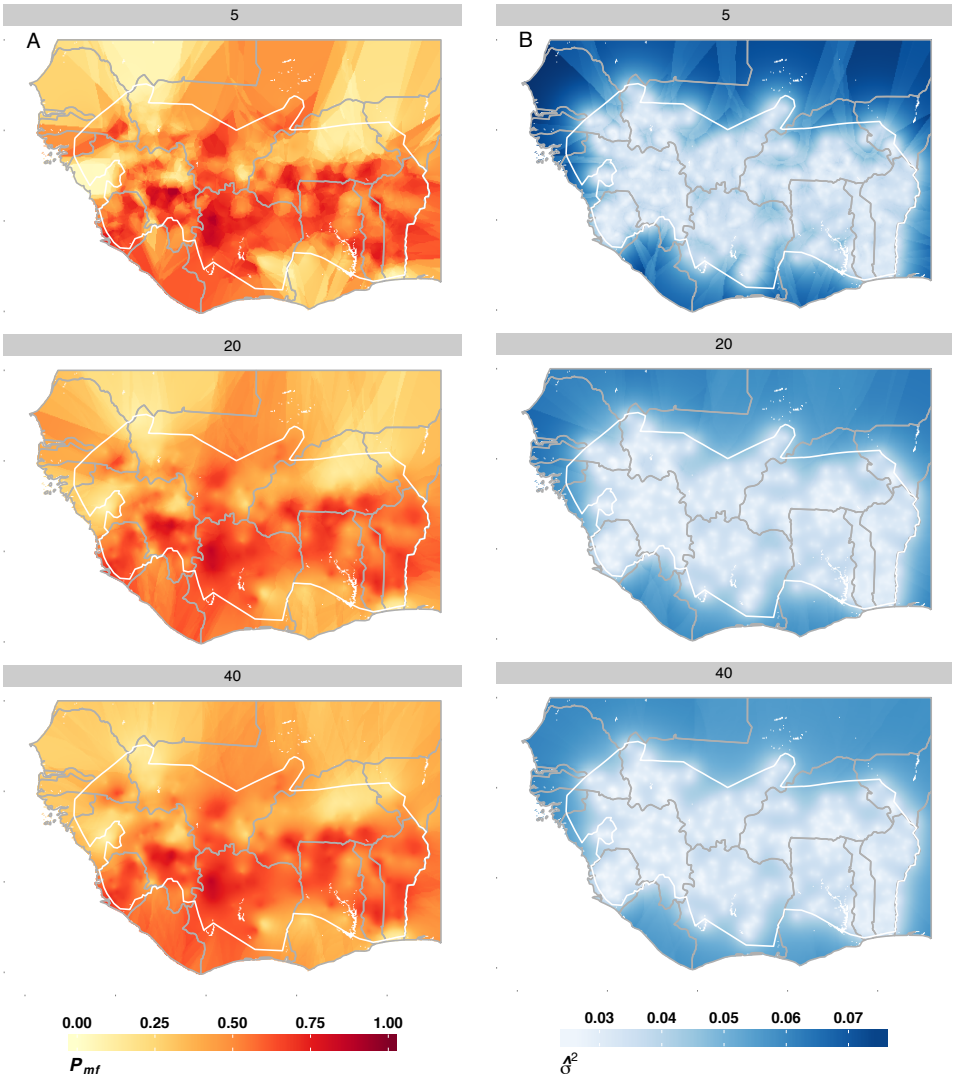


Figure 2-18: Ordinary kriging surfaces with locally estimated mean.
Left panels (A) are mean microfilarial. Right panels (B) are estimated variance of kriging predictor. The number in the panel titles above each map denotes the number of nearest neighbours used in estimating the local mean.

Finally, kriging incorporating an external trend (explanatory covariates of known onchocerciasis risk factors) was used to predict surface including a number of different external trends. The trend surfaces used are listed in **Table 2-2** as models A-H and described in the legend of **Figure 2-9**. The resultant maps are displayed in **Figure 2-19**. Panel **A** which assumes a stationary mean is the equivalent of simple kriging and resultant map is identical to **Figure 2-17 A**. Panels D, F and H include distance to nearest river as an explanatory covariate and the effect of this covariate is visible in the model output. Panel C includes a 2nd order polynomial trend on the x-y coordinates and takes negative values in the grey shaded areas.

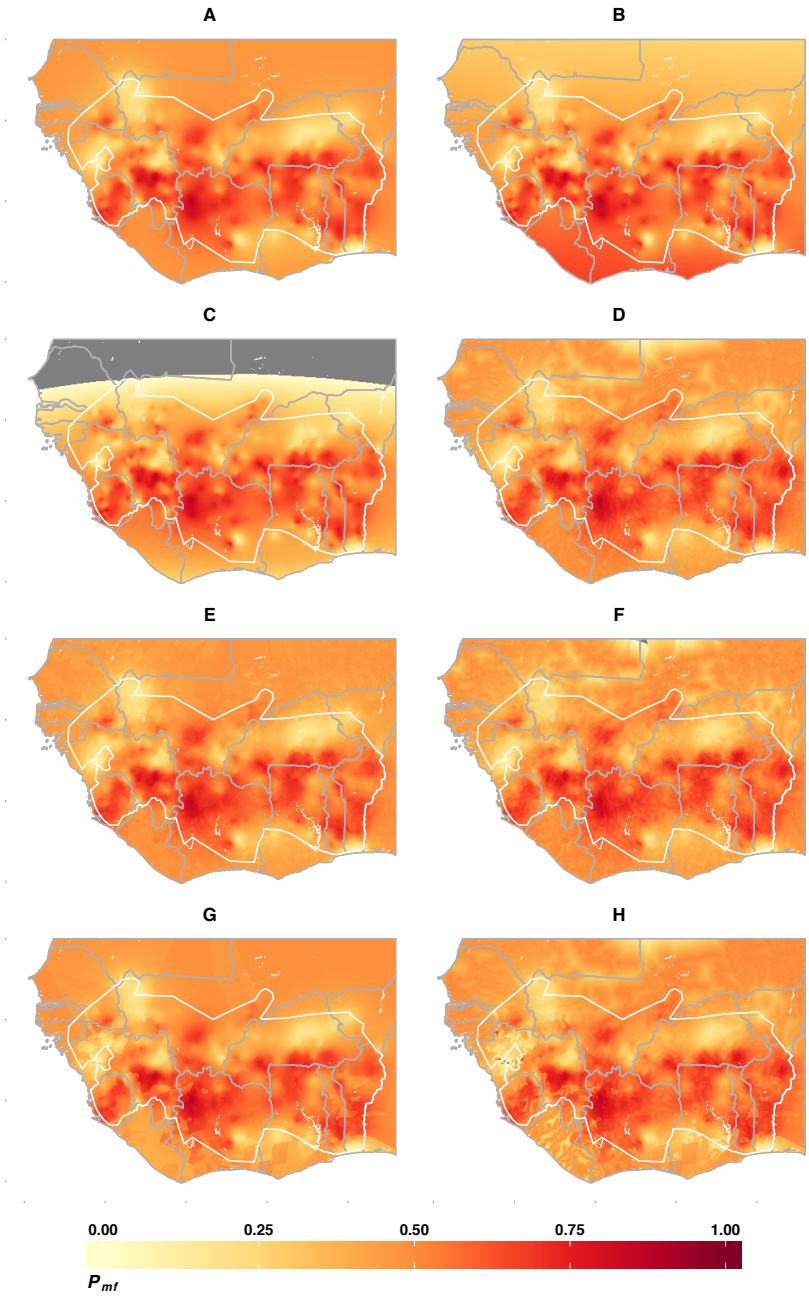


Figure 2-19: Maps of microfilarial prevalence estimated using external trend kriging.
Surfaces incorporating different trend models to account for a spatially varying mean were used to estimate microfilarial prevalence. The panel labels A-H refer to the trend model used listed in **Table 2-2** and described in **Figure 2-9**.

Cross validation of these kriging maps was carried out as described in **Chapter 3 Section 3.3.12**. The average linear correlation across all k-folds for randomly partitioned data is displayed in **Figure 2-20** for each of the trend surfaces modelled. There is good linear correlation between observed and estimated values across all trend surfaces. The points on each scatter plot are coloured by their

residual deviance, rescaled to take values between 0 (observed and estimated are equal) and 1 (the maximum residual deviance between observed and estimated) and give a visual indication of spread about the mean. **Figure 2-21** displays the same scatterplots for a cross validation exercise estimating prevalence in spatially clustered 'clumps' of *J*-nearest neighbours, with *J=74*. It is clear from the spread of values that there is a lower degree of linear correlation than in the randomly partitioned dataset. The values of the validation statistics for these cross validation exercises and an additional spatially clustered exercise using *J=30* nearest neighbours are given in **Table 2-7**. The best performing model overall was Model H which included a categorical variable for vector species and an interaction between distance to the river and the vector species.

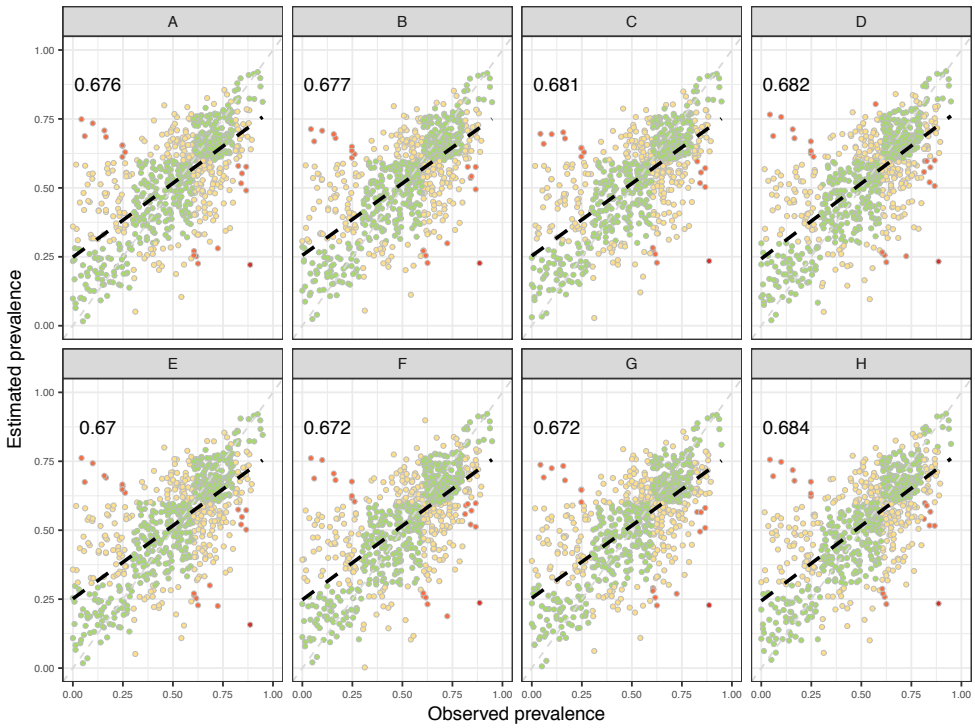


Figure 2-20: Linear correlation between observed and estimated mf prevalence using randomly partitioned validation data. Panels A-H represent models listed in **Table 2-2**. Green points denote a match between observed and estimated endemicity class according to the class boundaries given in **Chapter 1, Section 1.3.2** *Nodule Palpation and Rapid Epidemiological Mapping of Onchocerciasis*. Yellow denote a difference of one endemicity class between observed and estimated at those locations (e.g. a mesoendemic point is predicted to be hyper-endemic), orange and red points denote serious deviations of 2 or 3 endemicity classes respectively between observed and predicted endemicity classes (the most extreme example is a holoendemic point predicted to be hypoendemic and *vice versa*). The line of best fit is plotted as a black dashed line, and the 1:1 line of perfect fit is plotted as a faint grey-dashed line. The R^2 correlation coefficient is shown in the top right of each panel.

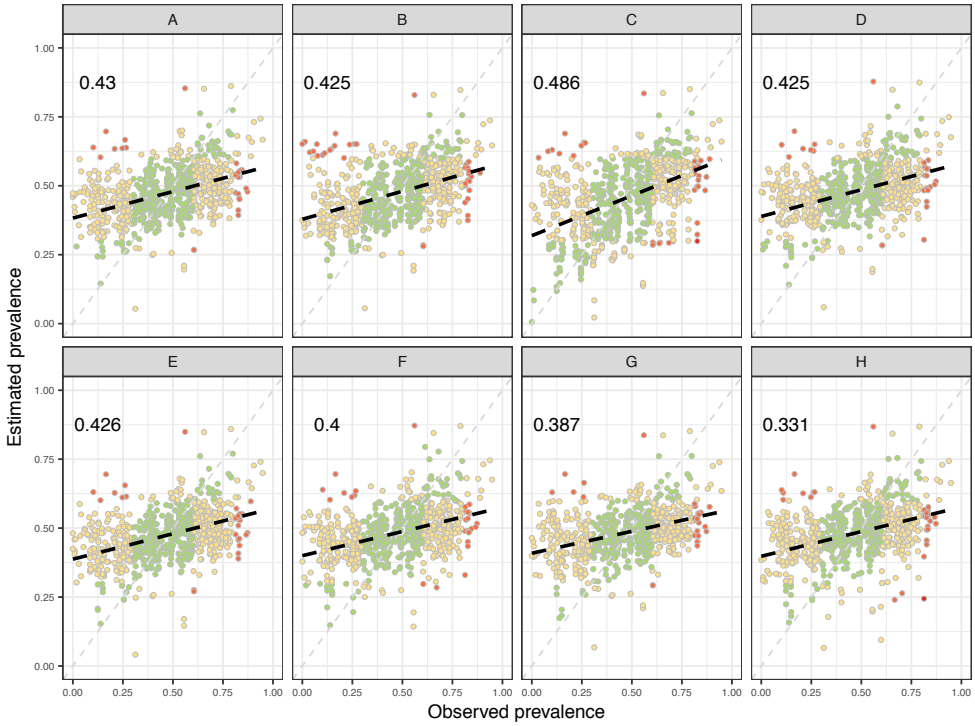


Figure 2-21: Linear correlation between observed and estimated mf prevalence using spatially clustered validation data. Panels A-H represent models listed in **Table 2-2**. Point colours and panel details are as **Figure 2.20**. The validation points in these panels are the *J=74* nearest neighbours of a randomly selected starting village.

Model		Spatially clustered *K*-folds *J*-NN = 74, *K* = 10			Spatially clustered *K*-folds *J*-NN = 30, *K* = 10			Random *K*-folds *N=74, K = 10*		
		Corr	**MAPE**	**MPE**	**Corr**	**MAPE**	**MPE**	**Corr**	**MAPE**	**MPE**
A	$\mu(mf) \sim 1$	0.367	0.156	-0.002	0.529	0.155	0.037	0.695	0.120	-0.004
B	$\mu(mf) \sim x + y$	0.389	0.153	0.001	0.493	0.156	0.036	0.694	0.120	-0.004
C	$\mu(mf) \sim x + y + x^2 + y^2 + xy$	0.422	0.151	0.019	0.522	0.157	0.034	0.700	0.119	-0.003
D	$\mu(mf) \sim Rd$	0.302	0.163	0.002	0.534	0.155	0.035	0.701	0.120	-0.003
E	$\mu(x) \sim Hr_i : i = 1,\ldots,6$	0.352	0.157	-0.007	0.522	0.156	0.038	0.692	0.120	-0.003
F	$\mu(x) \sim Hr_i + (Rd \cdot Hr_i) : i=1,\ldots,6$	0.259	0.165	-0.002	0.525	0.157	0.035	0.695	0.121	-0.003
G	$\mu(x) \sim V_s : s=1,\ldots,9$	0.390	0.152	-0.010	0.469	0.164	0.047	0.692	0.121	-0.003
H	$\mu(x) \sim V_s + (Rd \cdot V_s) : s=1,\ldots,9$	0.288	0.161	0.009	0.484	0.160	0.039	0.702	0.120	-0.003

Table 2-7: Validation statistics for kriging models incorporating an external trend.
J-NN refers to *J*-nearest neighbours, whilst *K* refers to the number of folds used in the validation exercise. *Corr* is the Pearson's correlation between observed and expected values, MAPE is the mean absolute prediction error (average magnitude of difference between observed and estimated) and MPE is the mean prediction error (mean difference between observed and estimated).

2.5 Conclusions

To the best of my knowledge, here presented is the first in-depth investigation into the existence of spatial autocorrelation in the Onchocerciasis Control Programme epidemiological dataset, for the purposes of estimating prevalence prior to the start of control interventions in West Africa.

Using a variety of spatial modelling techniques, it has been shown that there exists both a significant, positive degree of global spatial autocorrelation, as well as cluster of local spatial auto-correlation, most likely resulting from so called 'hotspots' of transmission (**Figure 2-10** and **Figure 2-11**). Because we can identify clusters of villages with similar endemicity levels, maps which incorporate information about local influences on the response variable will be more useful than those that don't. An important result from this Chapter is that models which simply describe the variation present in the available input data may not be as useful as those that use additional information to try to explain the unobserved, underlying data generating mechanism.

Multilevel B-splines are a highly robust, quick method of interpolation that use a summation of interpolations over a range of spatial scales. They are excellent when there are many data points, although when a K-folds cross validation was performed using 74 randomly selected points in each fold (with $K=10$), the Pearson's correlation between estimated and observed values was comparable, but not as high as that of any kriging model. When performing the cross validation using spatially aggregated K-folds datasets the performance drops from a correlation of 0.686 to 0.272. Only one out of the tested kriging models fails to achieve at least this degree of correlation between estimated and observed values.

I also found that the two best models in terms of highest correlation between estimated and observed values included a spatially varying mean trend, both of which included distance to the nearest river stream. This is an important result as it shows there could be a biological basis for including this covariate in models to explain the observed pattern of microfilarial prevalence. It also gives the justification for attempting to find a more sophisticated approach to modelling endemicity in the following chapter.

Choosing a spatial covariance function for modelling the data often appears to be a somewhat arbitrary choice. For this study, I decided against using the spherical model because it assumes zero spatial autocorrelation at separation distances above the range parameter, φ. Given that onchocerciasis is a vector-borne disease and that flight distances from vector breeding sites can be very large, in this setting I want to assume that although covariance approaches zero at large distances it does not vanish completely. The exponential model fitted this assumption and it is also the simplest function to model, having only one parameter. The decline in biting rates by blackfly vectors at different distances from the river has been shown to decline in an exponentially decaying fashion, so there is a biological plausibility to using the exponential function to model the spatial covariance between observed points of microfilarial prevalence.

I found a nugget effect to be an important component of the spatial autocorrelation present in the data. This is not wholly unexpected since from a practical point of view it can represent measurement error (i.e. the variance in the response variable at zero separation distance). Skin-snipping is known to be an imperfect detection method, particularly at low microfilarial densities and this result is concomitant with that observation. In conclusion, models incorporating spatial covariance and spatially varying mean trends may provide the most accurate realisation of the true unobserved underlying surface for onchocerciasis.

Chapter 3. Bayesian geostatistical modelling of the prevalence of human onchocerciasis using environmental covariates

A modified version of this chapter has been published: **O'Hanlon SJ**, Slater HC, Cheke RA, Boatin BA, Coffeng LE, et al. (2016) Model-Based Geostatistical Mapping of the Prevalence of Onchocerca volvulus in West Africa. *PLoS Negl Trop Dis* **10**: e0004328.

3.1 Chapter Aims And Objectives

The overall aim of this chapter is to improve the accuracy and utility of the prevalence distribution maps presented in Chapter 2. This will be achieved by meeting the following objectives: a) to identify a 'best' set of uncorrelated environmental covariates that can be used as explanatory covariates in a predictive model; b) to use the identified covariates and spatial autocorrelation in a Bayesian generalised linear spatial model (GLSM) to present a geostatistical map of human onchocerciasis (microfilarial) prevalence at the inception of the former OCP in West Africa; c) to map areas where high transmission occurred prior to control.

3.2 Introduction

In the previous chapter an attempt was presented to estimate the number of persons positive for microfilariae within the OCP area by using spatial interpolation and kriging. Although a good fit to the observed data may be obtained using these methods, they lack predictive power in un-sampled locations and make assumptions about the continuity of the outcome of interest between data locations.

In this chapter, the spatial modelling methods are greatly extended to introduce a GLSM which may incorporate one or more explanatory covariates, typically remotely-sensed environmental layers, to improve the predictive ability of the maps, particularly in areas where no ground truth data are available.

Given that modelling studies have shown that the feasibility of reaching, by 2020-2025, the operational elimination thresholds proposed by APOC (*182, 183*) depends on the level of initial onchocerciasis endemicity (*108, 109, 184*), production of this map will help to understand where these goals may be achieved with current intervention tools and where additional tools may be needed.

Currently, national control programmes and the WHO African Programme for Onchocerciasis Control (APOC) are shifting their focus from morbidity control to elimination of the infection through community-directed treatment with ivermectin (CDTI). Using 15–17 years of annual (or biannual) ivermectin distribution in the absence of vector control, elimination has been documented in some former OCP foci in Mali and Senegal (*185, 186*), and some APOC foci in Nigeria (*187*), providing proof-of-principle for the goals set by the WHO for elimination of onchocerciasis in selected African countries by the year 2020 (*188*). However, a key question that needs urgent investigation is where else will it be feasible to achieve the 2020 goals in the former OCP and APOC areas using current intervention strategies, and if not, by when these goals could be achieved, or what other (novel) strategies are required.

3.3 Methods

The methods stages are divided into 3 broad stages: 1) identification of the best set of uncorrelated potential explanatory environmental covariates, which may be used in the linear predictor of a geostatistical model; 2) determination of the most parsimonious model using one or more of the identified covariates and validation of model performance, and 3) generation of predictions and quantification of uncertainty at target locations in the study region.

3.3.1 Data description

For a description of the pre-control data locations and how they were selected see **Chapter 2 Section 2.3.2** *Description Of The Onchocerciasis Control Programme Epidemiological Database* and **Section 2.3.3** *Selection Of Pre-Control Data.*

3.3.2 Geographical limits of the mapping

In this chapter, the mapping was extended beyond the range of the OCP boundaries to see better the effect of making predictions in un-sampled locations, distant from known input data points. The mapping was restricted to the geographic extent visible in **Chapter 1**, **Figure 1-2**, which encompasses the OCP and immediate surrounding areas in West Africa. The OCP boundaries themselves were largely defined according to the primary objectives of eliminating blinding onchocerciasis (in savannah areas), and subsequently protecting the original area (of 654,000 km^2, spread over Burkina Faso, south-eastern Mali, south-western Niger, and northern Côte d'Ivoire, Ghana, Togo and Benin) from reinvasion by infective flies (*86, 99*), which led to the extensions listed in **Table 2-1** in Chapter 2. River basins in which the blackfly vectors of onchocerciasis can breed were identified by the OCP for the different species of the *Simulium damnosum* complex (*189, 190*).

The northerly limits of the mapping were allowed to extend to areas characterised by extremely arid, desert conditions where transmission of onchocerciasis is known not to occur. No pseudo-absence points were generated in these areas; the model was allowed to estimate prevalence values at target locations based on distance from the true input data locations and the values of environmental factors associated with the data. The easterly limit extends to the fringes of Nigeria, where control is under the auspices of APOC and estimation beyond these areas is not considered.

3.3.3 Climatic and environmental data

An ensemble of explanatory environmental and climatic covariates were identified, which were deemed to be useful in the estimation of microfilarial prevalence, building on the methodology presented in a previous mapping study for another filarial infection (*137*), namely, lymphatic filariasis. Initial inclusion was broad and partly based on knowledge of parasite and vector ecology (e.g. distances of villages to fast flowing rivers), as well as variables considered to affect their distribution and life-cycle (e.g. climatic data). Climatic data were extracted from WorldClim Global Climate Data, http://www.worldclim.org (*191*).

Hydrologically derived variables were extracted and calculated from the Hydro1K database (*192, 193*), and an alternative set of hydrologically derived variables was extracted from the Rivers of Africa dataset (derived from HydroSheds) (*152*) which includes alternative metadata about the stream lines.

Normalized Difference Vegetation Index (NDVI) data were extracted from the Global Inventory Modeling and Mapping Studies (GIMMS) dataset for the years 1981 to 1986 (*194-196*) (the earliest data available in this dataset).

Data layers, generated from the MODIS instrument on board the Terra satellites, were also used. These layers included land surface temperature (day and night), land cover classification, NDVI and enhanced vegetation index (EVI) (*197-200*). A layer of vegetation type based on physiognomy and floristic composition from White's Vegetation Map of Africa was also extracted and used (*201*).

Initially layers for altitude derived from digital elevation models (DEMs) from different mapping projects were tested. They were extracted from the Global 30 Arc-Second Elevation (GTOPO30) DEM (*202, 203*) and the Global Land One-km Base Elevation Project (GLOBE) (*204*). Where necessary, the original data sources were assembled together to form contiguous coverage of the study area and re-projected to the output map projection (Lambert Azimuthal Equal Area). Maps were then re-sampled to 25km^2 (the resolution of the output map), cropped to the borders depicted in **Chapter 1 Figure 1-2** and finally, those cells that overlay inland water bodies were masked as "no data" values.

Each pixel of the output map ($n = 112,295$) therefore had a value for each of the environmental covariates. The values of these covariates at each survey village were sampled from these layers. The initial ensemble of covariates tested and the data sources are summarised in **Table 3-1**. For further details of the pre-processing of these datasets and maps of the covariates that were included in the model, see **Appendix C** *Processing of environmental covariates*.

Dataset	Layers included	Pixel resolution	Coverage	Projection	Datum	Temporal coverage	Source	References
WorldClim	Min Temp,Max Temp,Mean Temp,Precipitation,Bioclim 1-19	1Km	Global	Geographic	WGS84	1950-2000	http://www.worldclim.org/bioclim	(*191*)
HYDRO1k Elevation Derivative Database	Slope,Aspect,Compound Topographic Index,Stream Lines[1]*,Drainage Basins,Elevation[2],Flow Direction,Flow Accumulation	1Km	Africa	Projected	Lambert Azimuthal equal-area	NA	https://lta.cr.usgs.gov/HYDRO1K	(*192, 193*)
Rivers of Africa (Derived from HydroSHEDS)	Stream Lines[1],Strahler Order,River Gradient,River Type	Shapefile	Africa	Geographic	WGS84	2000	http://www.fao.org/geonetwork/srv/en/metadata.show?id=37333	(*152*)
Global Inventory Modeling and Mapping Studies (GIMMS)	Normalised Difference Vegetation Index[3]	8Km	Global	Projected	Albers equal-area conic	1981-2006	http://www.landcover.org/data/gimms/index.shtml*	(*194-196*)
MODIS Vegetation Indices (MOD13A1)	250m 16 days Normalised Difference Vegetation Index, 250m 16 days Enhanced Vegetation Index	250m	Global	Projected	Sinusoidal	2000-Present	https://lpdaac.usgs.gov/products/modis_products_table/mod13q1	(*198*)
MODIS Land Surface Temperature & Emissivity (MOD11A2)	8-Day daytime 1km grid land surface temperature, 8-Day nighttime 1km grid land surface temperature	1Km	Global	Projected	Sinusoidal	2000-Present	https://lpdaac.usgs.gov/products/modis_products_table/mod11a2	(*199*)
MODIS Land Cover Type Yearly (MCD12Q1)	Land Cover Type 2 (UMD)	500m	Global	Projected	Sinusoidal	2001-2012	https://lpdaac.usgs.gov/products/modis_products_table/mcd12q1	(*197, 200*)
UNESCO (White's) Vegetation Map of Africa	Vegetation type map	30 arc seconds	Africa	Geographic	WGS84	1968-1983	http://geonetwork.grid.unep.ch/geonetwork/srv/en/metadata.show?uuid=cc2fbd96-ae4c-4593-910f-c103b1b51299	(*201*)
Global 30 Arc-Second Elevation (GTOPO30)	Altitude[2]	30 arc seconds	Global	Geographic	WGS84	1996-Present	http://webmap.ornl.gov/wcsdown/dataset.jsp?ds_id=10003	(*202, 203*)
Global Land One-km Base Elevation Project (GLOBE)	Altitude[2]	30 arc seconds	Global	Geographic	WGS84	NA	http://www.ngdc.noaa.gov/mgg/topo/globe.html	(*204*)
Gridded Population of the World (GPW V3.0)	Population density, Population count	5Km	Global	Geographic	WGS84	1990-Present	http://sedac.ciesin.columbia.edu/data/set/gpw-v3-population-density	(*205*)
Global Rural-Urban Mapping Project (GRUMP) V1.0	Rural-Urban Extents	Shapefile	Global	Geographic	WGS84	1990-Present	http://sedac.ciesin.columbia.edu/data/dataset/grump-v1-urban-extents	(*206*)

Table 3-1: Sources of spatially referenced data considered for use as explanatory covariates.
[1]Environmental covariate layers for Log_{10} distance to river and distance to river were derived from these vector shapefiles
[2]Multiple digital elevation models from these sources were tested and the best performing variables selected for further testing. See Appendix C for details.
[3]This dataset is currently unavailable from the Global Land Cover Facility website.

3.3.4 Temporal trends in the data

Although the baseline survey data used in this study were collected over a period of 16 years (1974–1990), it was not necessary to include the effect of time as these data were assumed to be representative of the pre-control endemic equilibrium situation at the beginning of each OCP phase and commencement of intervention by river basin. Time-varying changes in the environmental covariates were also treated as negligible. The WorldClim dataset uses interpolated data averaged over 50 years, whilst the MODIS data, representing 10-day means, were averaged by ecological quarter (1st: Dec–Feb, 2nd: Mar–May, 3rd: Jun–Aug, 4th: Sep–Nov) across the entire duration of available data at the time of analysis (February 2000–July 2013). There was no discernible non-linear trend in NDVI from MODIS vs. NDVI from the Advanced Very High Resolution Radiometer, which has a much earlier temporal coverage (from 1981) but a coarser spatial resolution (64km^2 pixels). Therefore the long-term climatic and environmental data were implicitly assumed to be representative of the average conditions present in 1975-1990 even if portions of the covariate data had been collected prior or subsequent to this date.

3.3.5 Spatial dependence in the data and model-based geostatistical analysis

A geostatistical model (used to develop a continuous map of pre-control microfilarial prevalence across the OCP area) is simply a Generalised Linear Mixed Model (GLMM) with a spatial random effect (GLSM) that accounts for the spatial correlation in the data and fixed effects that account for the non-spatial elements. If we assume that the number of individuals found positive for *O. volvulus* microfilariae at location x_i is Y_i out of the total number of people examined by skin snip at such location, n_i, then Y_i is a binomial random variable, $Y_i \sim Bin(n_i, p_i)$ where p_i is the proportion of individuals positive for skin microfilariae at location i.

An exploratory tool widely used in geostatistics is the semivariogram (*175*). Under the assumption of spatial dependence, there is a relationship between the variance of a pair of random variables (in this study microfilarial prevalence) and their separation distance, *d*. The empirical variogram represents the relationship between that separation distance, binned into intervals, and the mean semivariance of infection prevalence between all pairs of locations within each distance bin (**Figure 3-1 A-B**).

First, a theoretical variogram was fitted to an empirical variogram controlling for any environmental covariates included in the model (**Figure 3-1B**). An exponential covariance function was used and the parameters of the theoretical variogram were used to generate sensible prior values for the spatial component of the model (see **Section 3.3.8** *Prior Specification Of Model Parameters*).

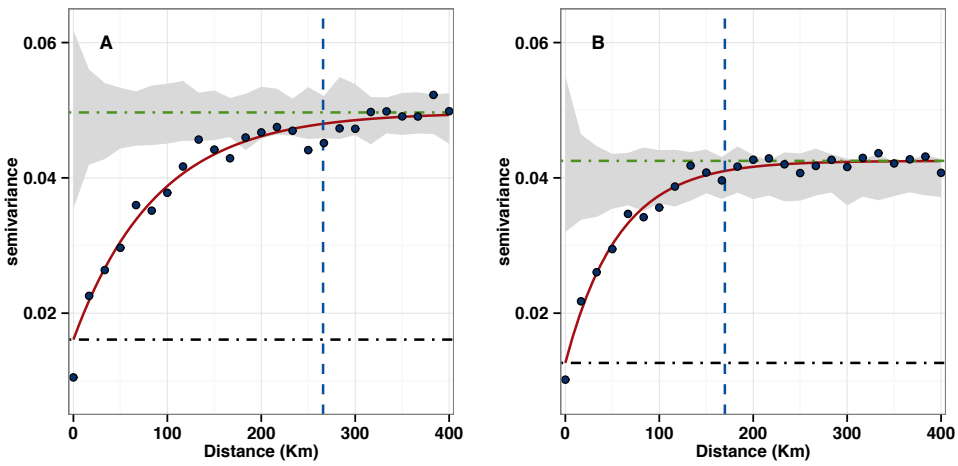


Figure 3-1: Variograms of spatial trends.
Panel A) The dark-blue dots are the empirical variograms assuming a constant mean trend across the study region (i.e. without the use of environmental covariates). Panel B) This variogram illustrates the spatial trend present in the data after accounting for the effects of explanatory environmental covariates by incorporating a spatially varying mean trend surface. The reduction in range, compared to the mean trend indicates that environmental covariates account for some of the observed spatial variation. The dash-dot green horizontal lines are the sill variance, σ^2 and are the estimated variance of the spatial process S(·). The dashed vertical lines indicate the range at which the variance asymptotically reaches 95% of its sill value and is given by $-ln(0.05)\phi$, where ϕ is the range parameter (see main text). Beyond this value, the covariance between locations does not depend on separation distance. The grey shaded areas represent the Monte Carlo envelope of variance expected by chance, estimated by randomly reassigning the values of prevalence at data locations and calculating the empirical semivariogram. Both variograms indicate significant spatial autocorrelation in the data.

Second, these values were used in a GLSM to fit a model of linear predictors (the environmental variables as fixed effects), conditional on some stochastic process $S(\cdot)$, i.e. the unobserved, underlying spatial process (which enters the model as a location-specific random offset in the linear predictor). Following (*130*), modifying the notation in (*137*), and conditional on $S(\cdot)$,the responses

Y_i/n_i , at locations x_i , for $i = 1,...,k$ (k = 737 survey villages) are mutually independent random variables whose conditional expectations, $\mu_i = E\left[\left(Y_i / n_i\right)\middle| S(\cdot)\right]$, are given by,

$$h(\mu_i) = S(x_i) + \sum_{j=1}^{p} \beta_j d_j(x_i) + U_i, \tag{3.1}$$

Where μ_i is the expected value of microfilarial prevalence at location x_i ; $S_1,...,S_k$ are the unobserved random variables for the $i = 1,...,k$ locations derived from the spatial process, $S(\cdot)$; $d_j(x_i)$ represents the value of the jth environmental variable at location x_i ; β_j is the regression coefficient of the jth variable, and the link function is denoted by $h(\cdot)$. For binomial data the link function is the logit function, i.e. $h(\mu) = \log\left[\mu/(1-\mu)\right]$. U_i are mutually independent Gaussian variables with mean 0 and variance τ^2 , which is known as the *nugget* variance. Such variables are sources of non-spatial error, e.g. measurement error from skin snipping.

The spatial signal $S(\cdot)$ has expectation $E[S(x)] = 0$, variance $Var[S(x)] = \sigma^2$ and correlation structure,

$$C(d) = Corr[S(x), S(x')] \tag{3.2}$$

Where $d = \|x - x'\|$ is the Euclidean distance between x and x' . The specification of this correlation function determines the smoothness of the underlying spatial surface (see **Section 3.3.8** *Prior Specification Of Model Parameters*). The vector of $j = 1,\ldots,p$ regression parameters is denoted as β and the parameters of the correlation function as θ .

This model specification and the data can be used to make predictions about locations of interest which may not have been sampled, but for which values of the regression coefficients are available (environmental covariates).

Continuing with the notation of (*130*), a vector of prediction may be specified for *target* locations $T=T(S^*)$, where $S = S(x_1^*), \ldots, S(x_m^*)$ are the values of $S(\cdot)$ at the x^* un-sampled locations where predictions are required (in this study, the 25km^2 grid covering the study area containing observed values of covariates at all m = 112,295 prediction locations). It is not possible to calculate the predictive distribution $[S^*|Y, \beta, \theta]$ directly, therefore Markov chain Monte Carlo (MCMC) simulation is used to generate samples from the conditional distribution, $[(S, S^*)|Y] = [S|Y][S^*|S, Y]$. Because S^* is conditionally independent of Y and β given S(·), values are generated by direct simulation from the conditional distribution $[S^* \mid S, \theta]$. (See (*130*) for further details and extensions to this framework.)

The modelling framework was implemented using the geostatistical modelling software from the package *geoRglm (Version 0.9-2)* (*179*) of the R statistical environment for computing (*Version 3.0.1*) (*178*).

3.3.6 Initial selection of environmental/climatic covariates

A systematic Bayesian step-wise procedure was used to select the best set of uncorrelated predictor covariates for use in the geostatistical model. This method builds upon that of Austin and Tu (*207*) by fitting univariate Bayesian GLSM models for both linear and quadratic functional forms of each of the initial predictor covariates. Model fit was tested by means of a *K*-folds cross validation, with $K = 1, \ldots, 10$. The data were randomly partitioned into a training set, consisting of approximately 90% of the data (N_{T_K} = 663), whilst the remaining 10% (N_{V_K} = 74) were held-out and used in the validation exercises. In each *K*-fold a different set of input data points was held out, such that each of the 737 data points appeared in one and only one of the *K*-fold validation data set. For this reason

$N_{V_{K=10}}$ comprised of 71 data points and $N_{T_{K=10}}$ 666 training data points (see **Figure 3-2** for locations of data points in each *K*-fold).

Each model was then independently fitted against the 10 training datasets, with priors for the spatial covariance structure re-estimated via fitting a theoretical variogram to the data each time. The Deviance Information Criterion (DIC), which is a Bayesian measure of model fit (*208*), was used to compare how well each covariate was able to estimate prevalence at a set of held-out validation locations. Ten datasets were chosen as a compromise between number of bootstraps and computational time.

For both linear and quadratic functional forms of all the input covariates, the DIC was computed for each bootstrapped model. For each covariate, the functional form yielding the average lowest DIC across all bootstraps was used to select either linear or quadratic forms for further evaluation. The covariates were then ranked by their resulting DIC values (lowest to highest) and an iterative selection procedure applied to select the best set of uncorrelated covariates. In the first iteration, the overall best-performing covariate (the covariate generating the lowest DIC) was selected and the correlation between values at all data locations between the model with this covariate and with all other covariates was computed.

Any covariates with a Pearson's correlation greater than 0.7 were deemed to be collinear and were eliminated from further model selection. The second iteration started with the next best remaining covariate, i.e. the covariate yielding the next lowest DIC remaining after the previous round. Collinear covariates were excluded as before and further iterations continued until all remaining covariates had either been eliminated or included in the set of uncorrelated covariates.

3.3.7 Building a multivariate model

Given the set of uncorrelated potential predictors identified, a standard Bayesian forward step-wise procedure was used to build the most parsimonious model that provided the best fit to the observed data. Selection was initiated by selecting the model which had the lowest possible DIC, identified from the previous step. Using this as a baseline model, bivariate models, each using an additional covariate, were run. Whichever, if any, of the new models had the greatest reduction in DIC from the baseline was selected to be the new baseline model, and additional covariates were added to this model.

At each step the covariate that effected the greatest marginal reduction in DIC was added. The step-wise procedure stopped when the addition of further covariates did not reduce (or increased) the DIC. **Table 3-3** lists the covariates that were included in the final Bayesian GLSM, their median posterior values and their 75% and 95% Bayesian credible intervals (BCI), respectively containing 75% and 95% of the posterior distribution for each parameter.

3.3.8 Prior specifications of model parameters

Initial fitting of semivariogram models suggested that the exponential correlation function gave an excellent fit to the data for both a mean trend without covariates (**Figure 3-1A**) and after controlling for the effect of environmental covariates (**Figure 3-1B**). The covariance $C(d)$ between a pair of variables (in this case microfilarial prevalence) located at points separated by distance *d* in the exponential model is given by,

$$C(d) = \exp\left(-|d| / \phi\right) \tag{3.3}$$

And the semivariance $\gamma(d)$ is given by,

$$\gamma(d) = \left[1 - \exp\left(-d / \phi\right)\right] \tag{3.4}$$

Where ϕ is the *range parameter*, determining the rate at which the dependence between $S(x)$ values decays with distance. Taking $-\ln(0.05)\phi$ gives the distance at which the semivariance reaches 95% of its maximum value (known as the sill or σ^2 following (*130*)). For each of the β_j environmental regression parameters, a prior distribution was assigned, such that $\beta_j \sim N\left(0, \sigma^2_{\beta_j}\right)$ and $\sigma^2_{\beta_j}$ was given a value of 10. Pragmatic priors for the spatial correlation function parameters (**eqn. 3.3**) were decided after initial investigation by fitting a theoretical variogram to the data conditional on the mean trend model specified by the environmental covariates (**Figure 3-1A**). The range parameter ϕ, was specified as a uniform prior across the interval $\phi' - 10 < \phi < \phi' + 10$ defined with 1000 discrete support points, where ϕ' is the maximum likelihood estimate of ϕ (used to find an initial sensible starting value), and the upper and lower limits of support were set to ensure exploration of a wide prior parameter space.

3.3.9 Markov Chain Monte-Carlo Chain Mixing

Although initial testing indicated that adequate mixing of most MCMC chains was achieved with a burn-in of less than 1.5×10^5 iterations, some bootstrap datasets did not converge, so the burn-in length was increased to 1.5×10^6 iterations to ensure that chains reached the posterior equilibrium distribution. A further 1.5×10^6 iterations were run to sample from the posterior distribution with thinning set to store every 500th subsample to reduce autocorrelation between successive subsamples, giving a total of 2000 samples from the posterior distribution. **Figure 3-3** depicts the traceplots for 3 independent MCMC chains of the ϕ parameter, which had different starting values, and **Figure 3-4** depicts the traceplots with independent chains for all fixed-effects model parameters (β_j) in the final predictive model.

3.3.10 Model Output

Model outputs consist of a series of realisations from the posterior distribution of both model parameters (**Figure 3-5**) and microfilarial prevalence for each map pixel. These may be summarised in a number of ways, e.g. a point estimate for each pixel was taken as the mean of the posterior distribution (**Figure 3-6**); pixels were discretized into 10% prevalence bands (**Figure 3-7**) and broad endemicity classes (**Figure 3-8**) following the endemicity classification for onchocerciasis used by the OCP (*34*).

The mean and median values of the posterior distributions were plotted on a histogram to examine the relative distribution of values from these summary statistics (**Figure 3-9**). It is also straightforward to derive quantities that may be of operational interest in control programmes, such as threshold probability maps (**Figure 3-10** and **Figure 3-11A**) or probability of endemicity class classification (**Figure 3-11B**).

A map of the predictive probability that prevalence did not exceed 35% was generated. The hypoendemic threshold of 35% is an operationally important threshold, where the burden of onchocerciasis as a public health problem is relatively low. The predictive probability that prevalence exceeds the epidemiologically important threshold of hyperendemicity (60%) was also mapped. At this threshold, there is a non-linear increase in morbidity experienced by the population.

To map the predictive probability that prevalence does not exceed 35%, approximations were taken from the predictive posterior at each target location (x^*) by calculating the proportion of realisations (r) from the predictive posterior that did not exceed the threshold,

$$(1/r)\sum_{i=1}^{r} P(0.35 < x_i^*) \qquad (3.5)$$

At each prediction location x^*, P takes a value of 0 or 1 depending on whether the threshold was exceeded or not. This surface is depicted in **Figure 3-10A**. The continuous probability surface that prevalence exceeds 60% is shown in **Figure 3-10B**, and a discretized surface of probability that

prevalence exceeds 60% is depicted in **Figure 3-11A**, classed into $P < 50\%$, $50\% \leq P < 75\%$ and $P \geq 75\%$.

A probability surface of endemicity class membership was also calculated for each pixel (**Figure 3-11B**). The most probable endemicity class was that within which the greatest proportion of realisations from the posterior distribution fell, with cut-offs set as follows (*34*): hypoendemic onchocerciasis: <35%; mesoendemic: ≥35% and < 60%; hyperendemic: ≥ 60% and < 80%, and highly hyperendemic (or holoendemic): ≥ 80%. The latter category, although not specified in (*34*), has been used to define more accurately the relationship between the prevalence of infection and onchocerciasis-associated morbidity, and to investigate the feasibility of reaching operational elimination thresholds with ivermectin distribution alone (*109, 144*).

3.3.11 Handling uncertainty in model Outputs

A key advantage of the B-MBG approach is that it permits robust evaluation of the uncertainty associated with model outputs. Uncertainty may be directly inferred from the posterior distribution of the predictions, with greater uncertainty indicated by a larger variance of the posterior distribution at prediction locations. The standard deviation of the predictive posteriors and the population layer were used to calculate a population-weighted index of uncertainty (**Figure 3-12**) following (*209*). Surfaces of the standard deviation and variance of the posterior distribution at each map pixel were also calculated (**Figure 3-13**). Finally, a map of the signal to noise ratio was calculated and is depicted in **Figure 3-14**.

The ability of the final B-MBG GLSM model to estimate the prevalence of *O. volvulus* microfilarial infection prior to the inception of the OCP control interventions in the area covered by the programme was assessed, for comparability with published work, using validation criteria which have been applied in other mapping exercises using Bayesian inference to generate predictive distributions. (e.g. (*209, 210*)). A *K*-folds cross-validation exercise (*211*) was performed as previously described. In each

cross-validation run (CV), a single subsample was retained to use as the validation dataset whilst the other $K-1$ datasets were used for training the model. A CV with $K=10$ was used; the data were randomly split into 90% training data and 10% hold-out validation data in each *K*-fold. The locations of the validation points in each fold are shown in **Figure 3-2**. The default random number generator (RNG) in R (the *Mersenne-Twister* algorithm (*212*)) was used to generate indices to split the data. For each CV, posterior predictive distributions of microfilarial prevalence were generated at unobserved validation locations for comparison against the known values. Using *K*-folds validation, each data point was used once in the cross validation. A number of validation indicators were then calculated according to the methodology used by the Malaria Atlas Project (MAP; http://www.map.ox.ac.uk/).

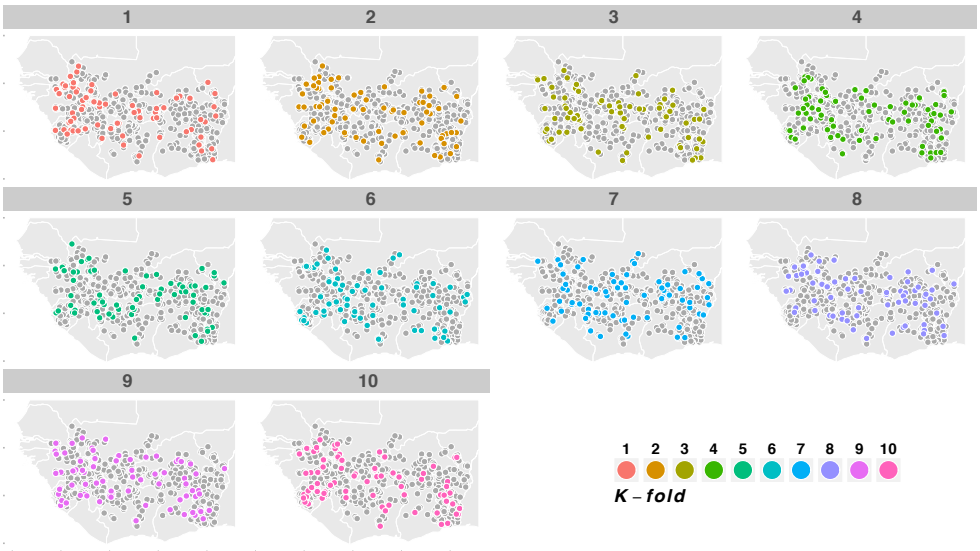


Figure 3-2: The location of held-out validation points used in each *K*-fold cross-validation dataset. The grey data-points represent the training data for each K-fold.

As a measure of the accuracy of point estimations the following statistics were calculated: (a) the overall Pearson-product-moment correlation coefficient (*r*) between observed and estimated values at validation locations (**Figure 3-15**); (b) the mean prediction error; (c) the mean absolute prediction error, and (d) the ability of the model to allocate endemicity class correctly (**Table 3-3**). The correlation coefficient provides a measure of linear association between observations and estimations and was also calculated separately for each *K*-fold cross validation. The mean prediction error is a

measure of overall bias, i.e. the tendency of the model to systematically over-estimate (if the value is positive), or under-estimate (if the value is negative) the true value. The mean absolute prediction error is a measure of overall accuracy, i.e. the magnitude of the difference between estimated and observed values (*213*).

Following (*209, 214*), measures of how closely posterior distributions matched their assumed error distribution were tested by means of a probability–probability (*P-P*) plot, which computes the expected value of microfilarial prevalence at each location, for a series of increasing probability thresholds (a total of 101 thresholds so that the range 0–100% can be explored), and compares them to the observed prevalence values (**Figure 3-16**). In a perfect error model, at a given probability threshold, the proportion of predicted values exceeding their observed value across all validation locations (the so-called 'coverage probability') should be equal to that probability threshold and therefore the *P-P* plot should show values lying perfectly along the 1:1 line. Deviations from this line indicate a tendency of the model to over- or under-represent uncertainty in point predictions. Comparisons were also made of the distribution of estimated and observed prevalence's at validation locations using a histogram (**Figure 3-17**).

3.4 Results

3.4.1 Model Parameters

As expected from the ecology of onchocerciasis, the $\log_{10}$-transformed distance to the closest river was a negative and statistically significant environmental predictor of onchocerciasis microfilarial prevalence (**Table 3-2**), confirming that as distance to rivers increases, microfilarial prevalence decreases. In addition to this, the only other (negatively associated) environmental covariates, which were statistically significant at the 95% highest posterior density (this is the shortest interval which contains 95% of the posterior probability), were: a) the second order polynomial term of log-transformed distance to river; b) the second order polynomial term of mean temperature of the wettest quarter (BIO8); c) precipitation seasonality (BIO15); and d) NDVI for the December–February ecological quarter. The spatial terms for the range (ϕ) and for the sill (or variance) of the spatial process (σ^2) were positive and significant predictors. At the 75% highest posterior density (HPD) (the interval which contains 75% of the posterior probability), in addition to the variables mentioned above, the intercept, the mean temperature annual range and the NDVI of the second (Mar–May) and third (Jun–Aug) ecological quarters were positively and statistically significantly associated (i.e. the 75% HPD interval for each parameters posterior distribution does not include 0) with microfilarial prevalence, whilst the second order polynomial terms of the temperature annual range and of the precipitation seasonality were negative and significant predictors.

The variograms depicted in **Figure 3-1** show that there was significant spatial dependence in the data (**Figure 3-1A**), even after accounting for the effect of environmental covariates (**Figure 3-1B**), as with increasing separation distance there was increasing semivariance lying outside the envelope of semivariance values expected by chance (the grey shaded area in **Figure 3-1 A-B** — a Monte Carlo envelope calculated from random reassignment of prevalence values to different data locations).

Variable	Median	95% BCI	75% BCI
Intercept	0.218	(–0.037 ,0.495)	(0.082 ,0.39)
BIO7: Mean temperature annual range	0.136	(–0.071 ,0.371)	(0.002 ,0.267)
BIO7: (Mean temperature annual range)^2	-0.087	(–0.22 ,0.053)	(–0.171 ,–0.008)
BIO8: Mean temperature of wettest quarter	-0.053	(–0.195 ,0.103)	(–0.151 ,0.027)
BIO8: (Mean temperature of wettest quarter)^2	-0.048	**(–0.1 ,–0.002)**	(–0.076 ,–0.018)
BIO15: Precipitation Seasonality	-0.455	**(–0.78 ,–0.143)**	(–0.638 ,–0.275)
BIO15: (Precipitation Seasonality)^2	-0.134	(–0.34 ,0.038)	(–0.243 ,–0.015)
BIO16: Precipitation of wettest quarter	0.013	(–0.192 ,0.24)	(–0.115 ,0.129)
Compound Topographic Index	0.014	(–0.061 ,0.109)	(–0.033 ,0.064)
Log_{10}(Distance to river+1)	-0.137	**(–0.215 ,–0.054)**	(–0.188 ,–0.094)
Log_{10}(Distance to river+1)^2	-0.034	**(–0.061 ,–0.007)**	(–0.051 ,–0.019)
NDVI (1st quarter: Dec–Feb)	-0.341	**(–0.637 ,–0.035)**	(–0.511 ,–0.162)
NDVI (2nd quarter: Mar–May)	0.201	(–0.204 ,0.561)	(0.001 ,0.445)
NDVI (3rd quarter Jun–Aug)	0.101	(–0.098 ,0.279)	(0.007 ,0.229)
NDVI (4th quarter Sep–Nov)	0.023	(–0.237 ,0.304)	(–0.15 ,0.167)
ϕ : range parameter	38.265	**(34.341 ,41.248)**	(36.303 ,40.187)
σ^2 : sill	0.790	**(0.707 ,0.874)**	(0.744 ,0.845)

Table 3-2: Posterior parameter values.
Posterior regression coefficient values and Bayesian credible intervals (BCI) for explanatory environmental covariates included in the linear predictor model and values for spatial covariance parameters. Parameters highlighted in bold and underlined were statistically significant (interval did not contain zero) at, the 95% BCI.

Plots of the highest posterior density (HPD) are shown in **Figure 3-3**, below. The HPD for some probability cut-off, *x%* is the shortest interval over the posterior distribution in which *x%* of the distribution falls. The HPD intervals were calculated using the package *LaplacesDemon Version 13.3.4* in R (3.0.1) (*215*).

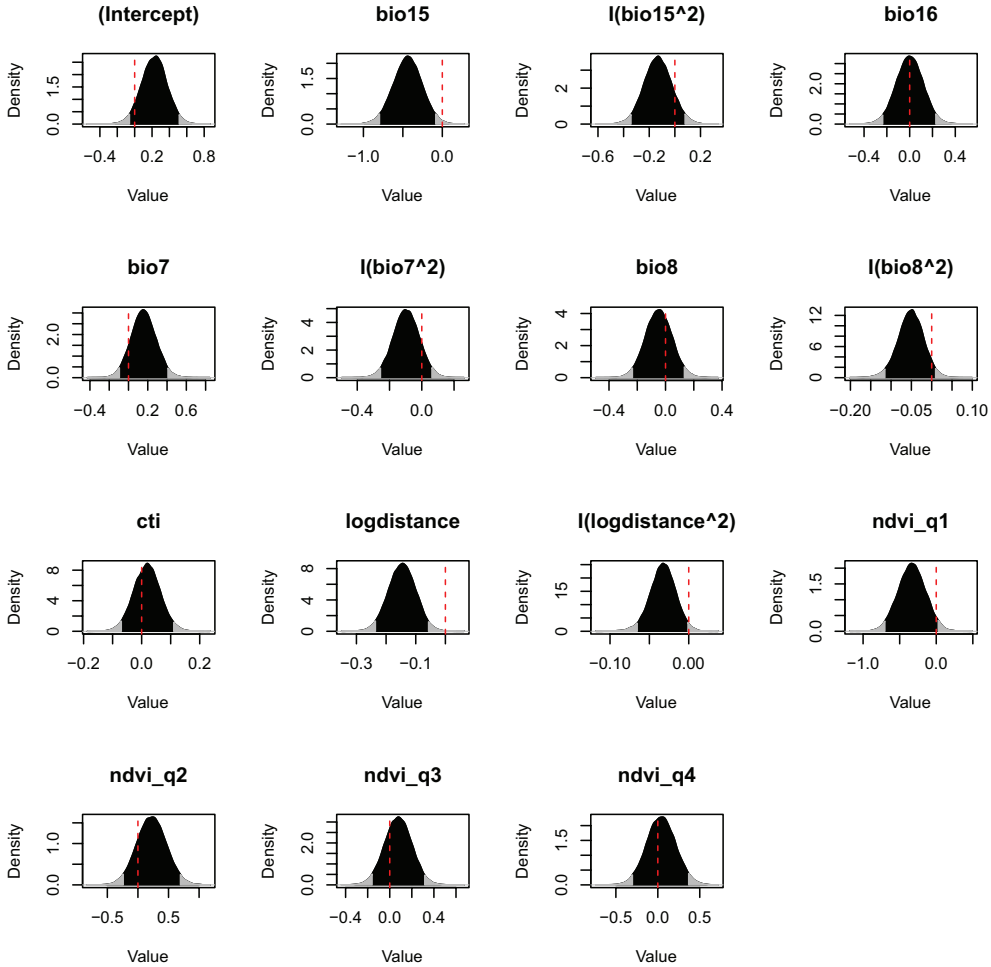


Figure 3-3: Highest posterior density plots for fixed effects.
The dark shaded regions contain 95% of the posterior distribution for each of the named β_j model parameters. The vertical red-dashed line crosses the x-axis at 0 (i.e. no effect). The model parameter *logdistance* is highly significant even at the 95% Bayesian credible interval.

3.4.2 Chain mixing and posterior convergence

Chains were initially run with different starting parameters to ensure good mixing of chains, and to estimate *a priori* the required length of burn-in period to be discarded. Since the time to convergence and thus the burn-in period is related to the initial set of parameters chosen and the parameters governing the step size of parameter is updated at each iteration of the model, once a suitable length for the burn-in period had been found, an arbitrarily long length of additional MCMC iterations was discarded to ensure that the numbers fell within the posterior distribution. Most parameters converged after just 5000 model iterations from very different starting values ($\beta_j = \{-5,1,5\}$). However, some parameters took longer to converge, taking $\sim 1\times10^5$ iterations for NVDI for the second ecological

quarter (see Figure 3.5 for chain labelled *NDVI_Q2*) and approximately 3×10^5 iterations for the range parameter ϕ (see **Figure 3-4**). Therefore, a long burn-in period was set of 1×10^6 iterations, which were subsequently discarded before sampling from the posterior distribution.

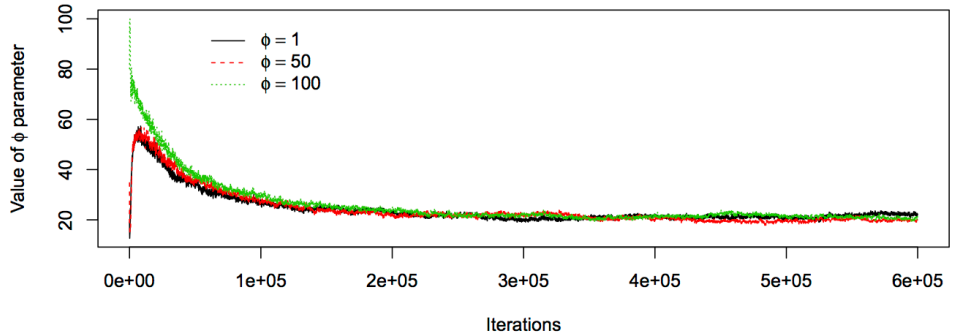


Figure 3-4: MCMC traceplot of the value of the range parameter.
Values of φ are plotted against iteration of the MCMC fitting process for three independent chains with starting values of φ=1, φ=50 and φ=100. Starting values were chosen to be far apart to test the robustness of the fitting process. Convergence was achieved after approximately $3\text{x}10^5$ iterations.

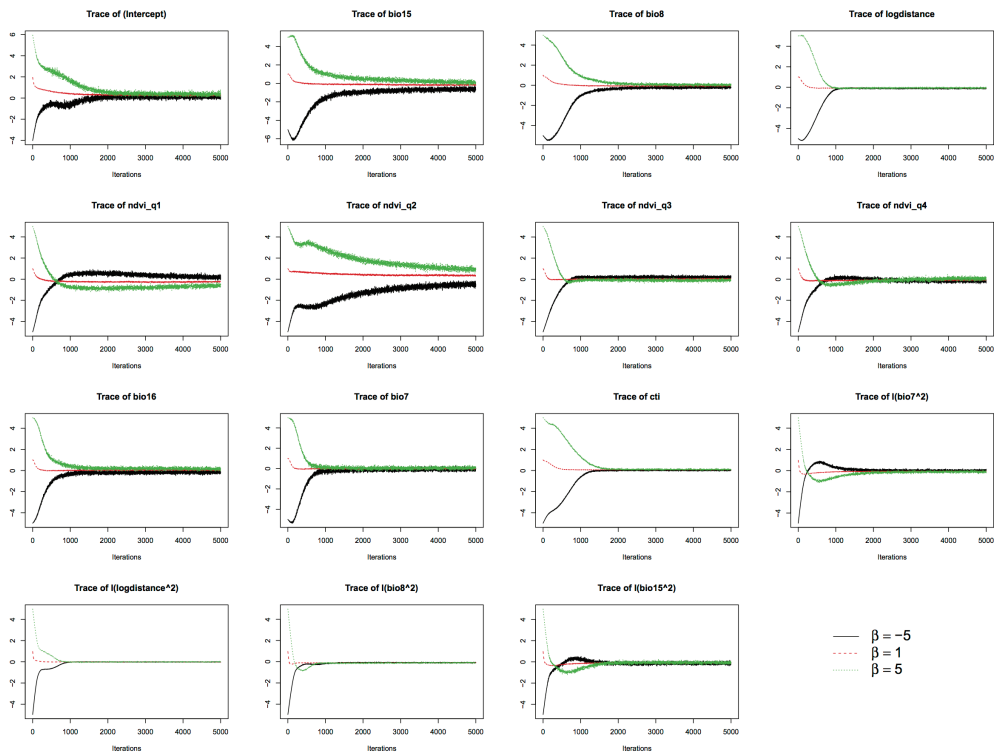


Figure 3-5: MCMC traceplots of all β_j regression coefficients.
The plots show the value of each regression coefficient (*y*-axes) at each iteration (*x*-axes) of the MCMC fitting process, for three independent chains with different starting values. Posterior parameter values are shown in **Table 3-2**.

3.4.3 Model output: map of baseline microfilarial prevalence prior to OCP interventions

The mean of the predictive posterior microfilarial prevalence in the population aged 5 years and older is shown in **Figure 3-6**. The mean *O. volvulus* prevalence in pixels within the boundaries of the OCP area was 45.2%, ranging between 2.3% and 90.0%. The mean prevalence in pixels outside the OCP control programme was 18.2%, ranging between 0.5% and 70.1%.

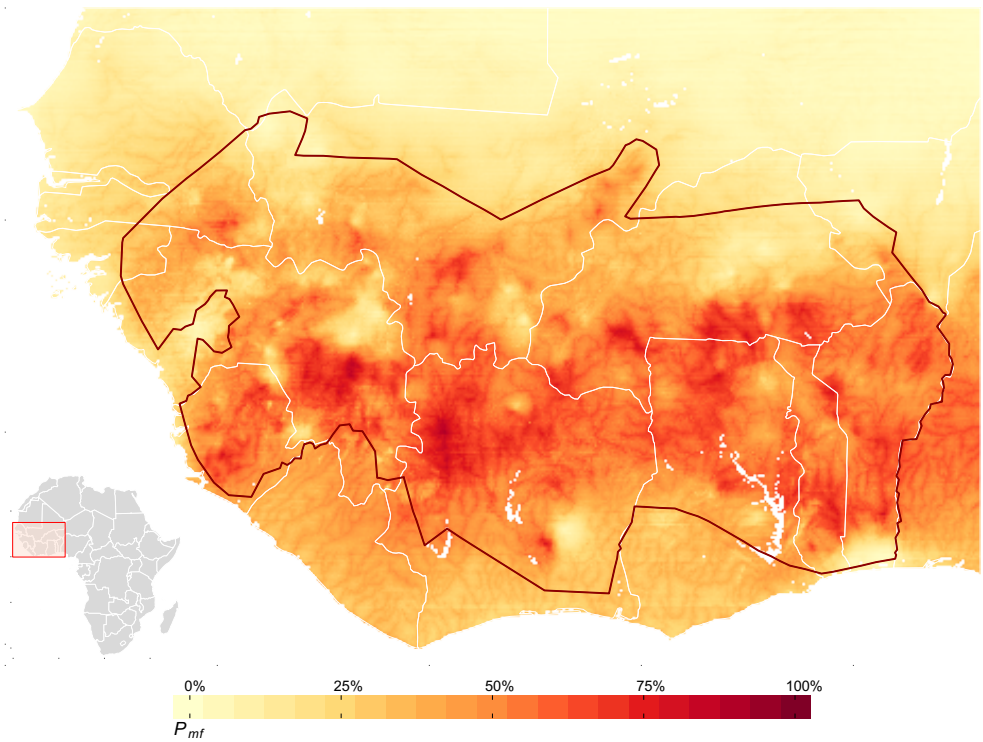


Figure 3-6: The pre-control spatial distribution of onchocerciasis microfilarial prevalence in west Africa.
The mean of the predictive posterior distribution of microfilarial prevalence in the population aged 5+ (denoted P_{mf} on the figure legend) for each pixel in the study area. The effect of the environmental covariate log(distance to river) is evident from the faint outline of the stream network appearing in this figure. The red boundary denotes the limit of the OCP area.

Figure 3-7 displays the mean point estimate binned into 10%-wide intervals of prevalence, making it is easier to see contiguous areas with similar prevalence values. However, this necessarily erodes some of the fine detail in the maps.

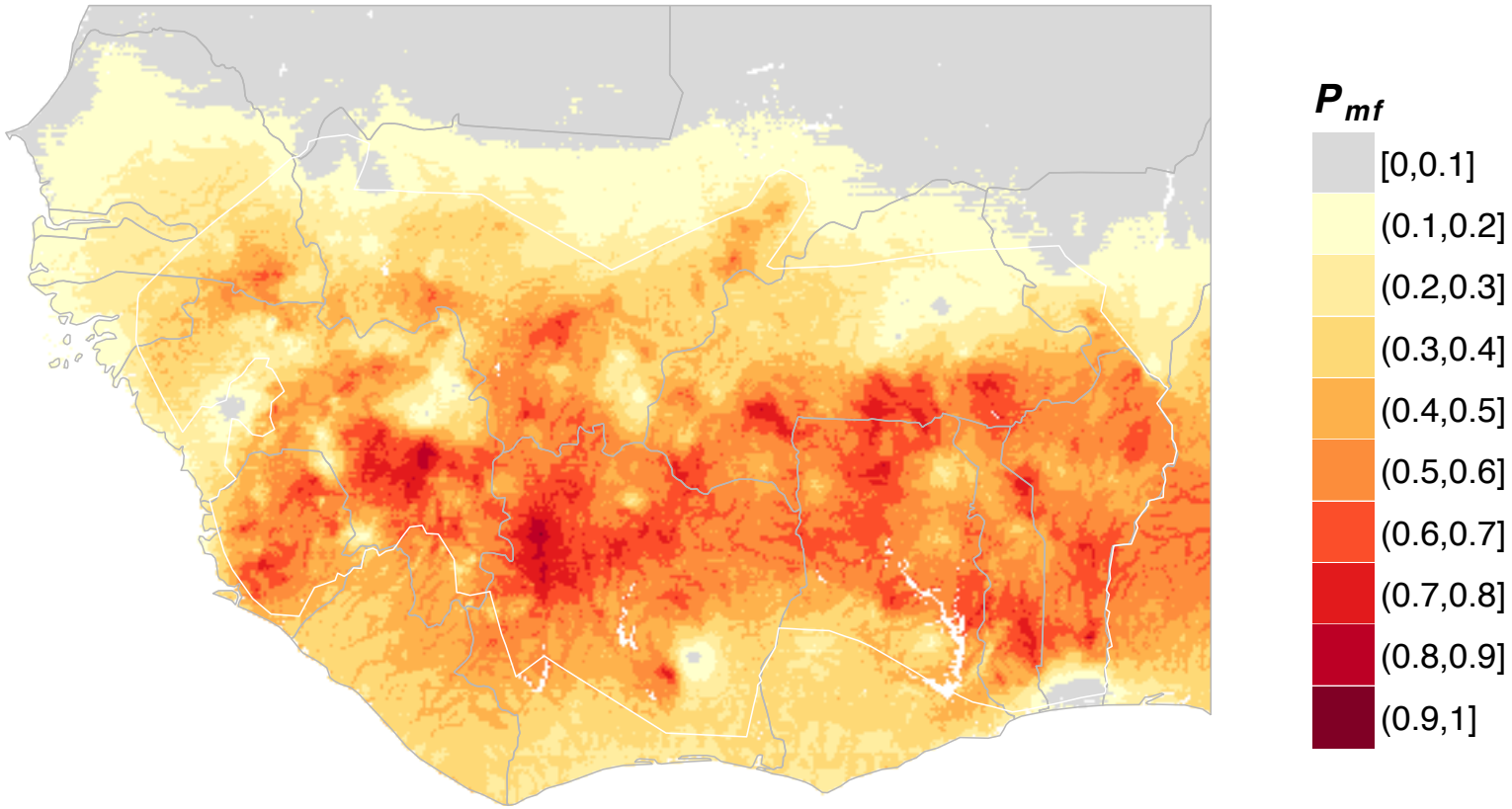


Figure 3-7: Alternative depiction of the pre-control distribution of onchocerciasis microfilarial prevalence in west Africa.
As in Figure 3.6, the pixels in this map are the mean of the predictive posterior distributions of microfilarial prevalence in the population, except that they have been binned into intervals of prevalence as shown in the key. There are no pseudo-absence data points in this model (pragmatically generated points of zero-prevalence in areas with known absence of disease) and therefore the grey areas in the north of the map, where predictions of prevalence are very low, are entirely determined by the value of the environmental covariates. The closest villages are too distal to exert a spatial effect. The white boundary denotes the limit of the OCP area.

A surface of mean prevalence discretized into the broad endemicity levels referred to in (*34*) and **Section 1.3.2** *Nodule Palpation And Rapid Epidemiological Mapping Of Onchocerciasis,* is shown in **Figure 3-8**, with the caveat that on this map only, the hypoendemic class has been sub-divided into two categories, adding a 0% to < 10% class, which is termed sporadic endemicity, to delineate areas with a low predicted prevalence. An endemicity class map of pre-control villages, used to generate these maps is displayed in **Section 2.2.3** *Selection of pre-control data* in **Figure 2-3**.

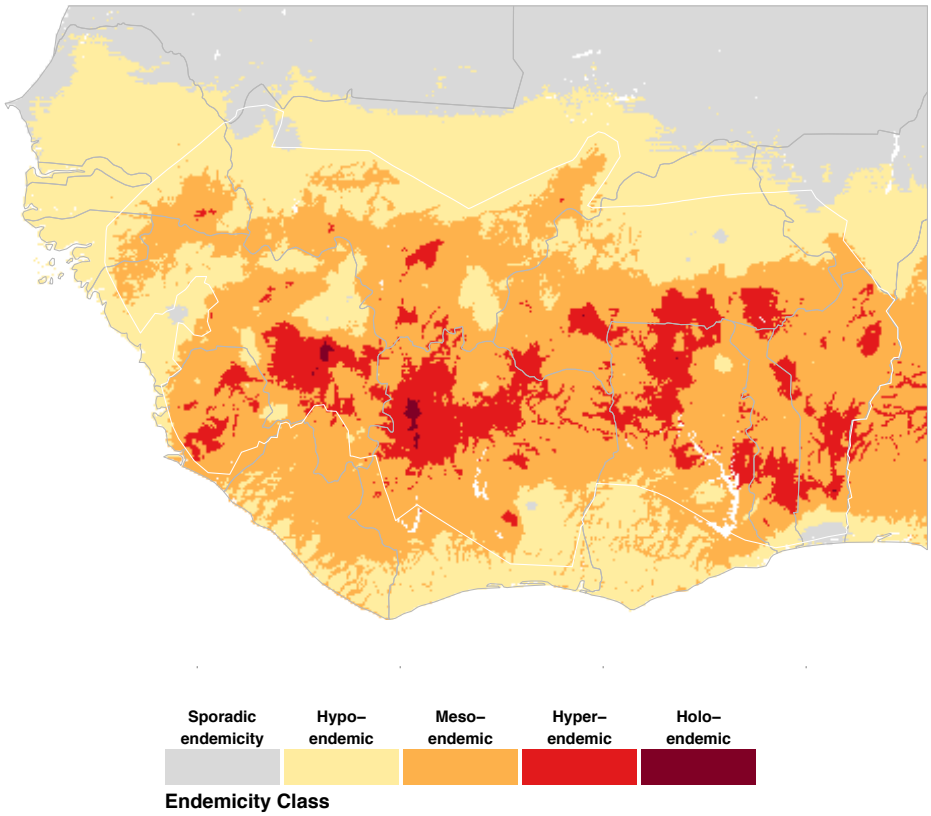


Figure 3-8: Endemicity class of microfilarial prevalence in west Africa in 1975.
Each pixel displays the endemicity class classifications based on the mean of the predictive posterior, based on OCP cut-offs, given in *Section 1.3 Diagnosing cases of onchocerciasis* with the exception that, in this map, to show areas where prevalence is very low a new non-endemic/sporadic endemicity class was added showing pixels where prevalence is <10%. Other classes are as follows: hypoendemic: >10% and <35%; mesoendemic: ≥35% and < 60%; hyperendemic: ≥ 60% and < 80%, and highly hyperendemic (or holoendemic): ≥ 80%.

A different summary of the posterior distribution could have been chosen to be displayed, e.g. the median of each distribution. Since a normal error distribution at each posterior predictive location was assumed, the values for the mean and median would ideally be identical. **Figure 3-9** is a histogram of

the values of the mean prevalence value (left panel) and median prevalence value (right panel) for all 112,295 prediction locations. These distributions are very similar, suggesting that the assumed normal distribution is a good fit for the posterior predictive distributions.

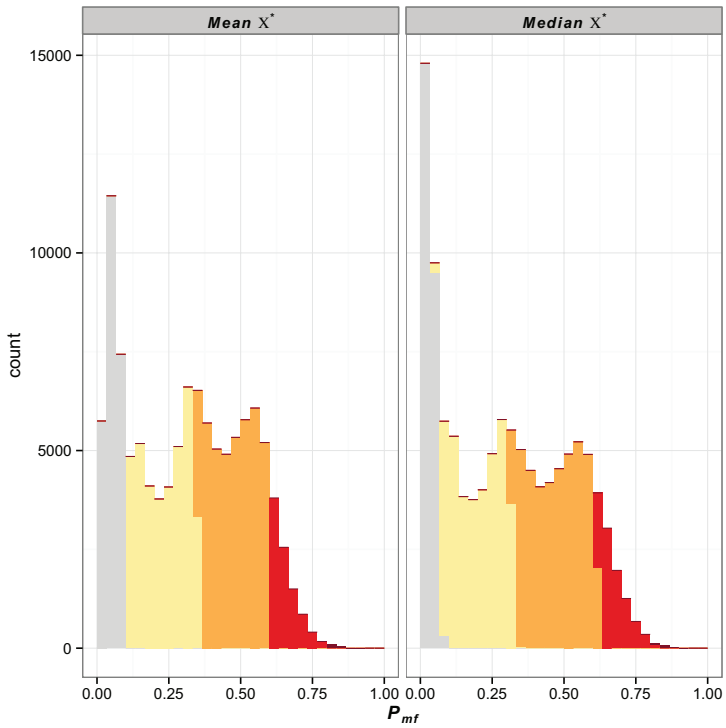


Figure 3-9: Histogram of mean and median microfilarial prevalence at all prediction locations.
Histograms displaying the frequency count (y-axis) of pixels in the study region, binned by prevalence (*x*-axis). Bars are coloured according to endemicity class (for key see **Figure 3-8**). Overlap in bars occurs when a prevalence bin crosses an endemicity class threshold. The similarity of the frequency distributions for both summary metrics suggests that the data are approximately normally distributed at each of the estimation locations.

3.4.4 Map accuracy and predictive probability maps

The predictive posterior distributions are an inherent realisation of the statistical uncertainty surrounding the mean estimates. The posterior distributions were translated into operationally useful maps, depicting areas where microfilarial prevalence is below or above a certain threshold. The probabilities that microfilarial prevalence is below 35% (where disease severity is low and reaching APOC's operational elimination thresholds should be feasible with the current intervention strategy of annual ivermectin distribution), or exceeds 60% (where disease severity is high and increased

ivermectin treatment frequency or additional intervention tools may be needed) are given in **Figure 3-10A** and **Figure 3-10B**, respectively. In both cases, the greatest levels of uncertainty were associated with areas where the posterior predictive mean of prevalence fell in the mesoendemic endemicity class.

Where predictions were either in hypo- or holoendemic classes, the probability of threshold exceedance was either very low or very high, respectively. The effect of the environmental covariate, distance to river, is visible in these continuous probability threshold maps. The outline of the stream network can be seen in **Figure 3-10A** as lighter shade in dark areas indicates a low probability that microfilarial prevalence is < 35% along the stream network. In **Figure 3-10B** the river network is also visible as darker colours in light areas, indicating a higher probability that microfilarial prevalence exceeds 60% close to the river.

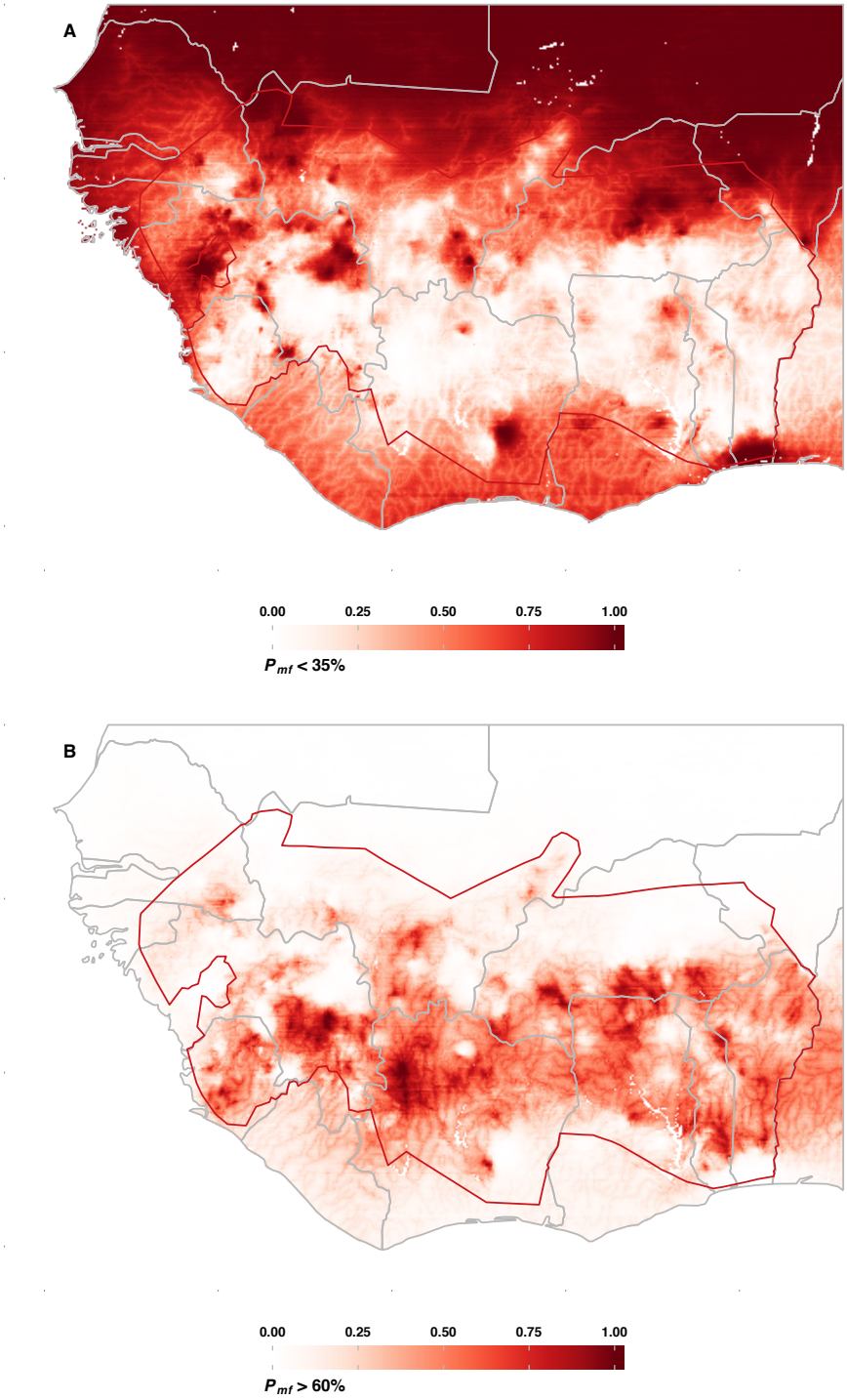


Figure 3-10: Maps of the probability of threshold exceedance.
A) Each pixel displays the posterior predictive probability that pixel microfilarial prevalence is less than the threshold for hypoendemicity (35%). B) Each pixel displays the posterior predictive probability that the pixel prevalence is greater than the threshold for hyperendemicity (60%).

Figure 3-11A also describes the probability that prevalence exceeds the clinically and epidemiologically important threshold of 60% microfilarial prevalence, binning pixels into classes of 0 to 50%, 50 to 75%, and 75 to 100% probability that areas are at least hyperendemic. Maps of this

nature will be extremely useful for the optimal targeting of current and novel control interventions as programmes strive to achieve morbidity control and infection elimination goals, because such maps facilitate visualization of contiguous areas where clinically and epidemiologically important thresholds are most likely to be exceeded, elimination goals may not be reached with current tools, and additional interventions or combinations thereof may be required.

There is a band of high-probability of exceedence of the hyperendemic threshold across the core area of the OCP, from Guinea, through Côte d'Ivoire, Ghana and southern Burkina Faso, to Togo and into Benin. This band also extends into Nigeria, despite there being no input data points there; a country which was later included in the mass treatment programmes under APOC (see Fig 2 and Fig 3 of (*216*) by way of validation of my results).

The average probability of endemicity class membership was 66% (**Figure 3-11B**), with a range of 34 to 100%. The most uncertainty surrounding endemicity class membership appears to be in the core areas of the OCP, predominantly in meso- and hyper-endemic classified pixels. However, examination of the signal-to-noise ratio map (see **Figure 3-14**, below) indicates that these are the areas with least dispersion in the posterior predictive mean. This indicates that the boundaries used to define endemicity classes are not well suited to this study. Pixels with estimates of prevalence close to endemicity class boundaries will be less easily classified to a given endemicity class, and instead a more finely discretized map, such as the map in **Figure 3-7** would be a better representation of the underlying true values.

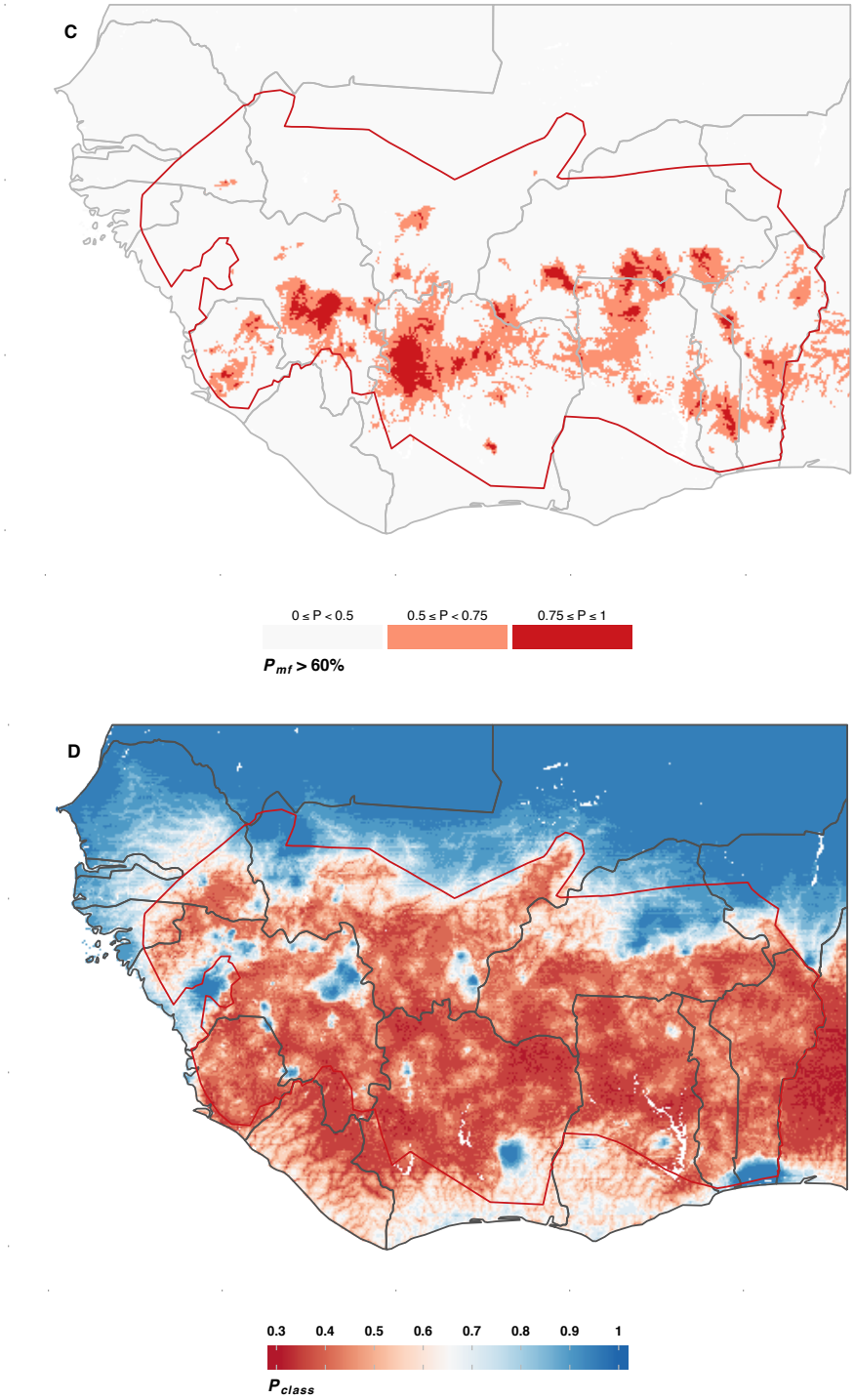


Figure 3-11: Probability of hyperendemic threshold exceedance and endemicity class membership.
Panel A) The data in this map are the same as in **Figure 3-10** but categorised into broad probability intervals, denoting the probability, P that P_{mf} exceeds 60%. Panel B) The probability of each map pixel being in the endemicity class to which it was assigned. With 4 endemicity classes the probability maps range from a minimum of 0.25 to 1, with a value of 1 meaning that all realisations from the predictive posterior were assigned to the same endemicity class.

Finally, summary statistics for the standard deviation (**Figure 3-12A**) and variance (**Figure 3-12B**) of posterior predictive distributions were computed. From comparing these maps with the map of input

data locations in **Figure 2-3**, the posterior predictive distributions are less diffuse (i.e. a better measure of central tendency) close to the input data locations. In the northernmost part of the map, in locations distant from input data locations, where environmental conditions are more extreme and arid, the standard deviation and variance also tend towards lower values.

The probable importance of the uncertainty surrounding endemicity class assignment was modelled as in (*209*), using a population-weighted index of uncertainty. Briefly, in areas where the probability of endemicity class assignment is not well defined (a low probability within the overall range of 0.25-1 in this study), and the population density is also high, the possible marginal effect of that uncertainty on the predictive posterior distribution of microfilarial prevalence will be high. The population-weighted uncertainty map depicted in **Figure 3-13** which has an uncertainty index value scaled between 0 and 1, is highly overdispersed, with > 90% of the pixels in the map having an uncertainty index of < 0.5 assigned to them.

The signal to noise ratio (SNR) is more often connected with the communication and electronics industries, but also has applications in interpreting biological data. The SNR indicates the ratio of signal strength (mean microfilarial prevalence) to noise (dispersion of the posterior distribution, as measured by the standard deviation). Pixels with a greater dispersion in the posterior distribution have a lower SNR, particularly in areas where the posterior distribution is highly over-dispersed and the estimated prevalence is low.

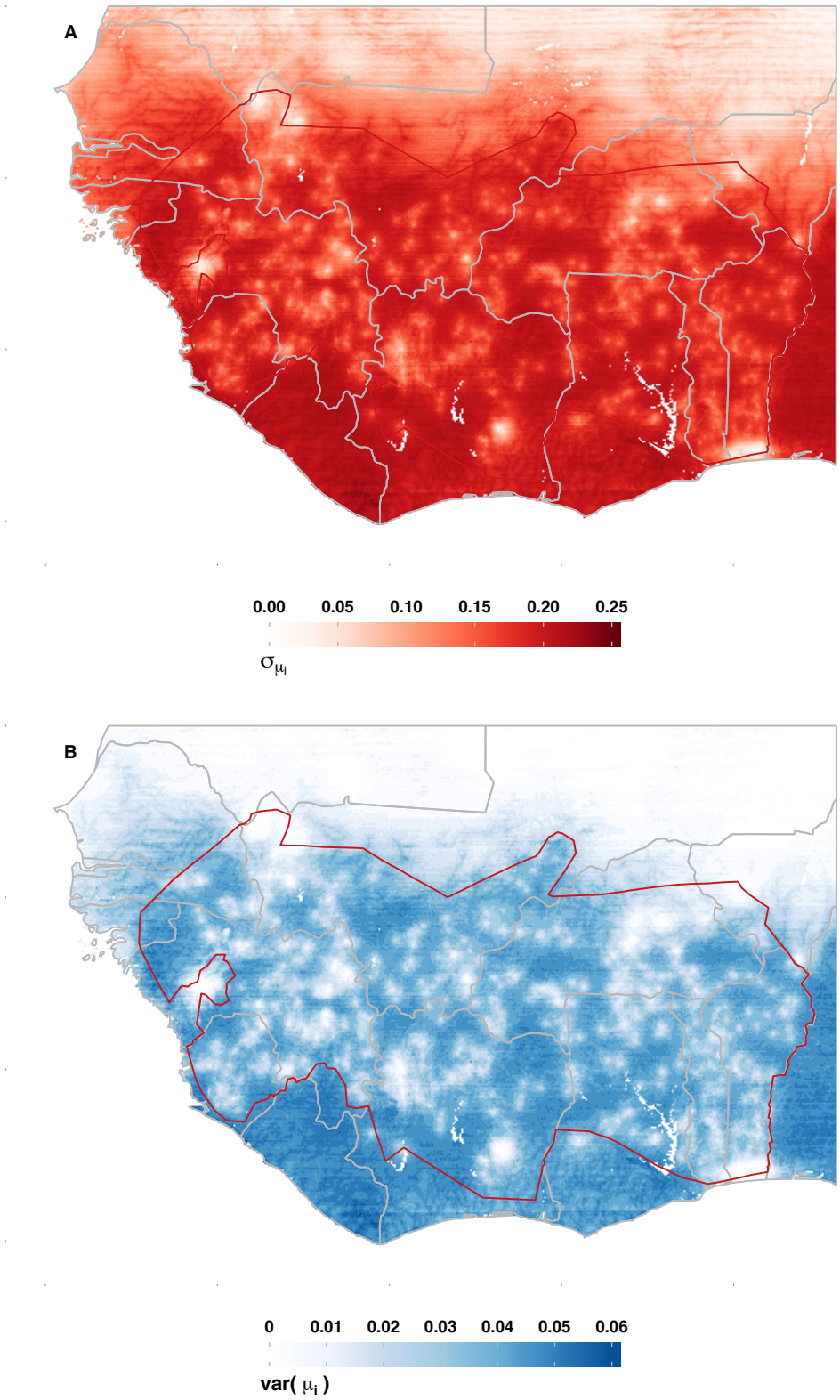


Figure 3-12: Maps of the standard deviation and variance of predictive posteriors.
Panel A) Standard deviations of predictive posterior distributions. Panel B) Variance of predictive posterior distributions. Higher values mean more diffuse posterior distributions and hence greater uncertainty in model estimates. The lowest uncertainty is found in pixels closest to data locations.

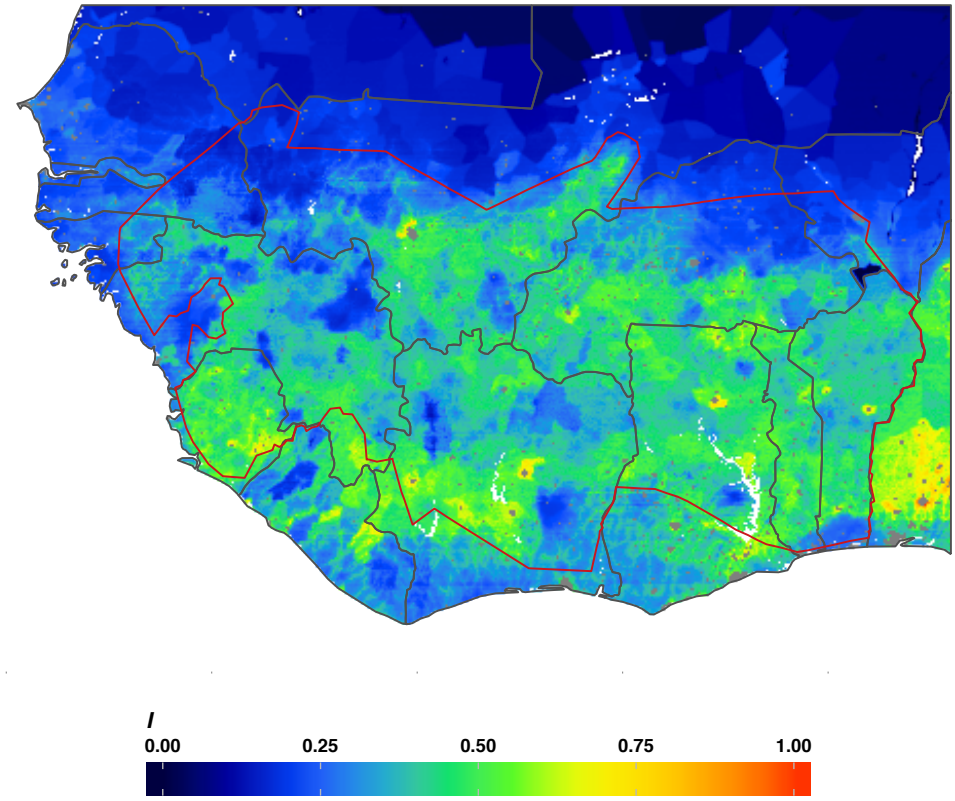


Figure 3-13: Population weighted index of uncertainty.
The population weighted index of uncertainty is calculated as in (*209*) and is a pragmatic representation of how important uncertainty in the predictive posteriors is likely to be. It is calculated by taking the $\log_{10}(pop_{75}+1)\times 1/P_{\text{class}}$ where pop_{75} is the population count for 1975 and P_{class} is the probability of endemicity class assignment.

In **Figure 3-14** the highest SNR values are within the OCP area, as is expected from a geostatistical model that has a dependence on Euclidean distance. In northernmost areas of the map, the SNR drops below 1 (above this threshold there is more signal than noise), due to the unavailability of ground-truth data-points nearby, despite the fact that prevalence is predicted to be low in these areas. Values below 1 indicate a degree of over-dispersion in the distribution; however, the low absolute values for the standard deviation or variance of the predictive posteriors in these pixels, indicates that the over-dispersion is skewed towards low values.

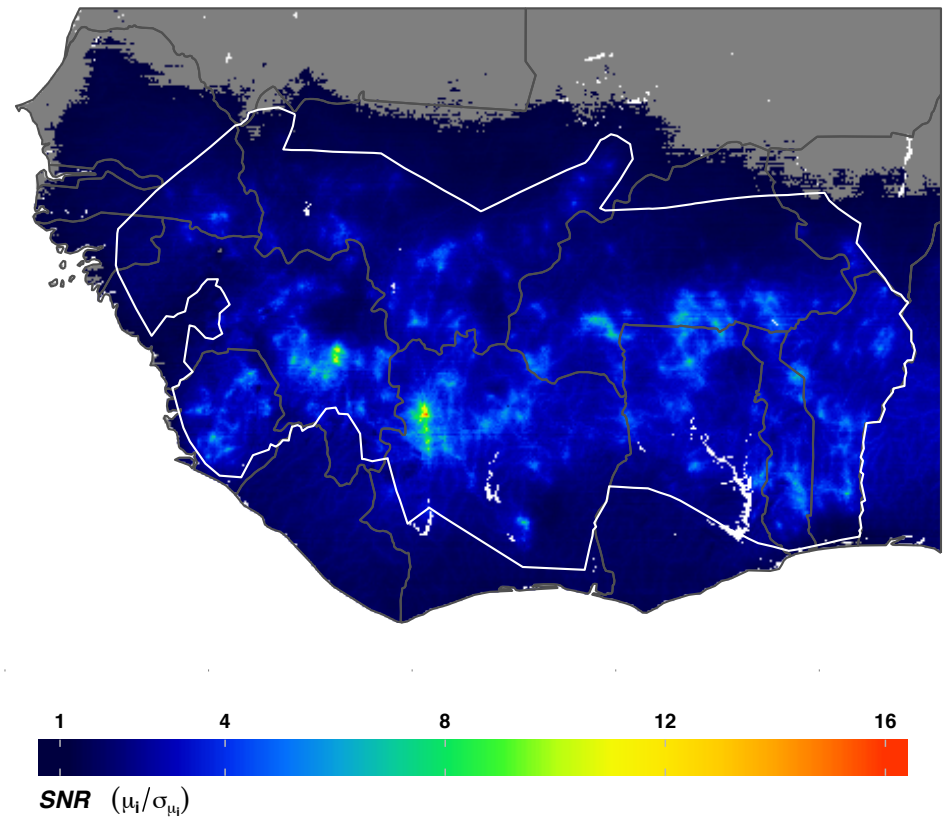


Figure 3-14: Signal to noise ratio of mean and standard deviation of predictive posteriors.
The strongest signal is found within the OCP borders, particularly in meso- and hyperendemic areas. A small cluster of very high prevalence villages between the Regions of Bafing and Denguélè in Côte d'Ivoire mean that all realisations from the predictive posterior are very similar, resulting in a very high SNR in that area (green and yellow area in Western Côte d'Ivoire). Grey areas in the north have a SNR of less than 1, indicating less signal than noise.

3.4.5 Model Validation

The overall correlation coefficient, r, between observed and estimated values at validation locations was 0.693 (**Figure 3-15** below). Calculating separately a correlation coefficient, ρ for each of the 10 K-folds individually, a range of correlation coefficients between 0.566 and 0.806, with a median of 0.699 was obtained. The mean prediction error was 0.77% (i.e., overall, observed prevalence was greater than the estimated prevalence by less than 1%). The mean absolute error indicated that, on average, there was a 12.3% difference between the observed and estimated microfilarial prevalence at validation locations. The linear correlation depicted in **Figure 3-15** is coloured by the mean absolute prediction error, rescaled to values between 0 and 1, with colours at the red end of the spectrum indicating a better fit of the estimated data to the observed data.

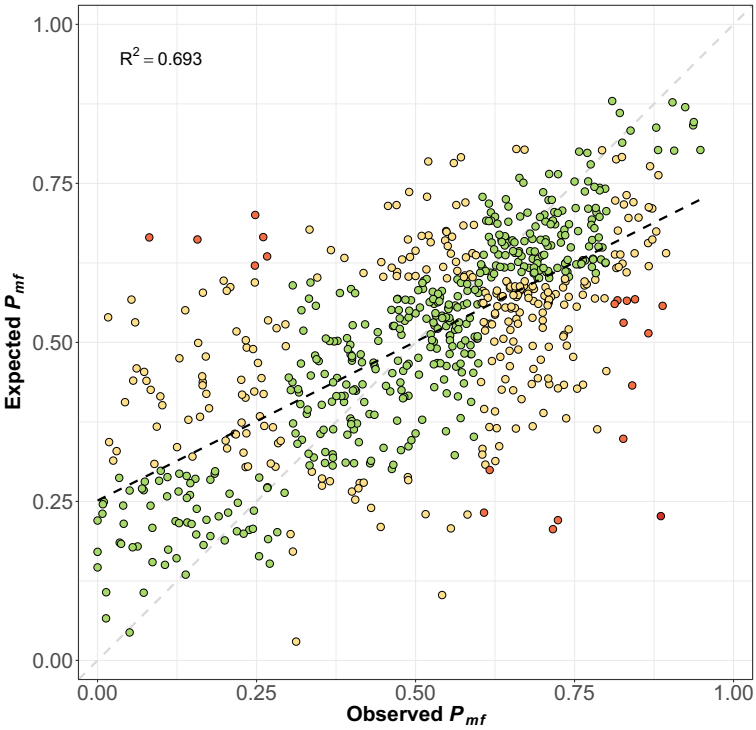


Figure 3-15: Scatter-plot of observed versus estimated microfilarial prevalence (P_{mf}) at validation locations. The linear correlation (Pearson's) between observed and estimated values was 0.693. Green points denote a match between observed and estimated endemicity class according to the class boundaries given in **Chapter 1, Section 1.3.2** *Nodule Palpation and Rapid Epidemiological Mapping of Onchocerciasis*. Yellow denote a difference of one endemicity class between observed and estimated at those locations (e.g. a mesoendemic point is predicted to be hyper-endemic), orange and red points denote serious deviations of 2 or 3 endemicity classes respectively between observed and predicted endemicity classes (the most extreme example is a holoendemic point predicted to be hypoendemic and *vice versa*). The line of best fit is plotted as a black dashed line, and the 1:1 line of perfect fit is plotted as a faint grey-dashed line. The R^2 correlation coefficient is shown in the top right of the panel.

The expected onchocerciasis endemicity class matched the observed endemicity class for 406/737 (52.7%) validation locations and was within one endemicity class at 711/737 (96.5%) of the locations, showing good agreement between observed and estimated classes. Serious misclassification of endemicity class (i.e. classifying a point to a non-adjacent class) occurred in 26/737 (3.5%) of the cases. The contingency table for the four endemicity classes in this study is given below in **Table 3-3**.

Expected Class	Observed Class				Total
	Hypo	Meso	Hyper	Holo	
Hypo	**91**	**32**	**9**	**1**	133
Meso	**74**	**172**	**123**	**9**	378
Hyper	**7**	**45**	**134**	**29**	215
Holo	**0**	**0**	**2**	**9**	11
Total	172	249	268	48	737

Table 3-3: Observed versus estimated endemicity class classification table.
Contingency table for the classification of survey villages according to observed and expected (model-derived) endemicity class for all validation locations. Serious prediction errors (*n=26/737*) occur in orange and red shaded cells, where the predicted class was more than one class classification away from its true class (e.g. a hyperendemic location classified as hypoendemic). Green cells denote where the predicted class matched the observed endemicity class.

Following (*209, 214*), a *P-P* plot was constructed to test how closely posterior distributions matched their assumed normal error distribution (**Figure 3-16**). The line is close to the 1:1 line of a perfect error model (dashed grey), but falls below the 1:1 line for values between probability thresholds of 0.15 to 0.60, meaning that there is a slight bias of the model to overestimate the probability of taking values closer to the middle of the range of possible prevalence values, whilst slightly under-representing extreme (very low or very high) microfilarial prevalence values.

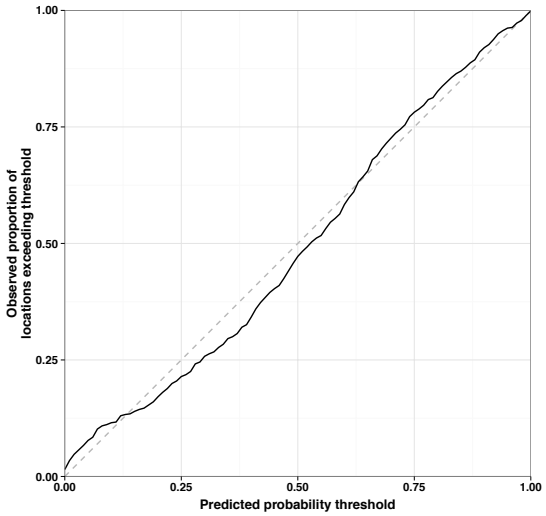


Figure 3-16: Probability-probability plot.
The proportion of true values of P_{mf} at validation locations that exceed the predicted probability threshold for a series of probability thresholds from 1% to 100%. The 1:1 line (dashed grey) in this plot represents a perfect model of uncertainty.

The phenomenon detailed in the previous paragraph is evident from the histogram in **Figure 3-17**. The histogram of observed prevalence at validation data locations is negatively skewed (biased towards higher values), whilst the histogram of estimated prevalence at the same locations exhibits a shift towards normality, with a greater proportion of estimated values falling in the meso- and hyper-endemic categories.

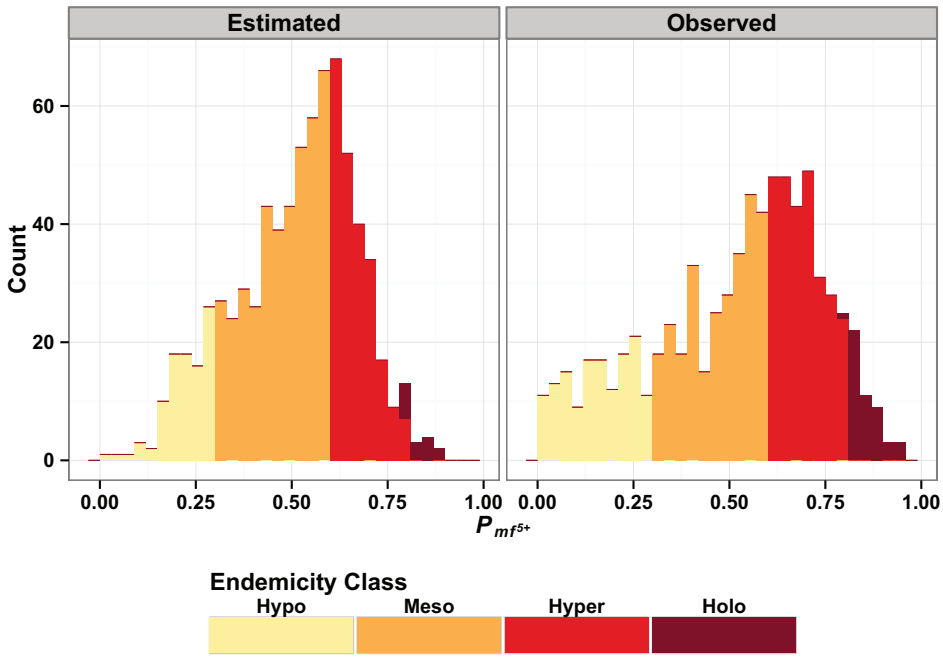


Figure 3-17: Frequency histogram of observed and estimated P_{mf} at validation locations.
Values are classified according to their true class (Observed P_{mf} panel) and the estimated endemicity class (Estimated P_{mf} panel).

3.5 Discussion

To the best of current knowledge, here presented is the first continuous map of the pre-control prevalence of *O. volvulus* microfilariae in West Africa, in the region previously covered by the OCP. In addition, environmental covariates have been identified that aid in the estimation of onchocerciasis prevalence in un-sampled areas and a robust handling of the uncertainty surrounding those estimates has also been treated. According to the model, the mean baseline infection prevalence in the OCP was 45%. Interestingly, Coffeng *et al.* (*125*) estimated that prior to the commencement of APOC, 45% of people, on average, were infected with *O. volvulus* in APOC areas.

Although levels of infection have decreased as a result of the interventions implemented by the OCP (*10, 86, 88*), and environmental change has occurred in the intervening years since 1975 (*217*), a map of the baseline levels of prevalence and endemicity provides an indicator of initial conditions for transmission, which are influential determinants of the feasibility of onchocerciasis elimination with ivermectin (*109, 140*). This is particularly relevant, since after cessation of OCP vector control operations in 2002, and subsequently in special intervention zones (*218*), blackfly vector populations have recuperated in many areas and there is evidence of continuing transmission within parts of the OCP area (*113*).

3.5.1 Representativeness of the data

The OCP epidemiological dataset represents the most complete source of standardised parasitological survey data for onchocerciasis in West Africa. However, although extensive surveys were conducted (see **Chapter 2.3.2 Figure 2-1** for a map of all locations surveyed during the OCP), only a small fraction of the total population was sampled (the subset of the OCP dataset which was selected in this study as the most representative of pre-control prevalence, consisted of results from skin snips of approximately 145,000 persons).

One of the problems faced, when using the OCP data for mapping purposes, was that the population and epidemiological surveys were not performed at random, introducing some bias in the data towards areas of higher endemicity. Using cartographic and entomological knowledge about the location of rivers, flow speeds and vector breeding sites, a village was classified as first-line if there were no other settlements between itself and the vectors' breeding sites (*219*). All first-line villages were proposed to be surveyed, with further surveys in a selection of villages conducted at regular intervals up- and down-stream from the breeding site, and a random sample taken from second- and third-line villages at increasing distance from the river (*220*). It was not necessary at the time to conduct a complete census of the resident population (although a census was carried out in each survey village) in the OCP area and its extensions, because control initially relied on vector control (and not on mass distribution of ivermectin, for which censuses are mandatory).

3.5.2 Use of environmental covariates and a spatial trend

The use of environmental covariates to improve model accuracy was investigated and a number of variables associated, or deemed to be potentially associated, with onchocerciasis was investigated. A model incorporating covariates was found to be a better predictor of onchocerciasis prevalence and endemicity than an intercept-only model, which incorporated only the spatial trend. Inclusion of covariates improved (reduced) the DIC values in model validation and improved the correlation between observed and predicted prevalence values at validation locations. In addition to the environmental covariates listed in **Table 3-1** in **Chapter 3.3.3** *Climatic And Environmental Data*, differences for the effects of variable geographical distributions of different members of the *S. damnosum* complex were tested but, surprisingly (given that members of the complex are the vectors), no significant influence of vector taxon on the model's predictions could be detected.

The spatial range value of $\phi \approx 38$ corresponds to a spatial autocorrelation that reaches 95% of its maximum semi-variance over a distance of approximately 113Km. The range parameter is consistent

with the observation that blackfly vectors can travel several hundred kilometres from a breeding site in flight (*150, 221-224*). However, Renz *et al*. (1987) found biting density to drop to 10% of its original level at distances of just 5Km from riverine breeding sites in a Sudan-savanna focus of Northern Cameroon (*151*). Other studies have, however, shown that people may be bitten by female blackflies in locations distant from any known breeding site (*225-227*) and that re-invasions of parous flies into areas previously cleared through larviciding can happen from breeding sites as far as 300Km away (*221*). This lends some epidemiological and ecological support to the operation of a spatial trend operating over this distance.

The methods of Slater *et al*. (*137*) were extended to carry out a robust, Bayesian step-wise selection procedure to identify the best set of uncorrelated explanatory covariates. The chosen covariates significantly improved the ability of the model to explain the variation in recorded microfilarial prevalence. The covariates also helped to define natural geographical limits to the mapping which would otherwise have been impossible, given the bias towards surveying villages with a high probability of having onchocerciasis. In other mapping studies, pseudo-absence points have been introduced into the data in un-sampled areas which are suspected to be non-endemic, to help to delineate risk-free areas. In this study this practice was avoided, and instead the environmental predictors, and the spatial covariance function of the model were used to define the limits. The model estimated very low prevalence in the very dry arid areas north of the OCP boundaries, without recourse to including pseudo-absence points.

In this study, microfilarial prevalence was statistically significantly and negatively associated with increasing distance from river streams in both the linear term (P-value < 0.01) and the quadratic term (P-value < 0.05). Distance between village and river is an important determinant of vector biting rate (*151*) and onchocerciasis endemicity when the river contains rapids and other suitable and identifiable vector breeding sites, according to a classification of first, second and third line villages (*32, 228*).

Mean temperature of the wettest quarter (BIO8) was also significantly and negatively correlated with prevalence (P-value < 0.05). Interpretation of the meaning of this is rather more difficult because the

wettest quarter is necessarily different across different map pixels, but it could suggest that conditions for blackfly breeding and transmission may be more favourable in areas in which the majority of precipitation occurs in the cooler parts of the year. This is linked with precipitation seasonality (BIO15), which was also significantly and negatively correlated with infection prevalence. Areas with greater intra-annual variability in rainfall (and therefore possibly representative of more seasonal transmission) were negatively associated with prevalence (the coefficient of variation of precipitation seasonality is the standard deviation of the weekly precipitation estimates expressed as a percentage of the mean of those estimates). The interpretation is that where transmission is interrupted in the dry season (e.g. the far north away from perennial rivers) it is hotter generally. This is consistent with the observation that there is a negative correlation with mean temperature of the wettest quarter. A higher prevalence is estimated where there is continuous transmission and lower where it is interrupted in the dry season.

Diawara et al (2009) reported the results of a study on the feasibility of elimination on onchocerciasis with current tools, i.e. ivermectin (*185*). Despite the fact that these foci contained hyperendemic villages, the foci are situated in areas of seasonal transmission. Thus whilst there can be great cause for optimism at the proof-of-principle of elimination using ivermectin in these three foci (*186*) it remains to be seen how applicable these results can be to other areas where transmission is perennial.

Areas with a more constant weekly rate of precipitation (perhaps indicative of perennial transmission), which also occurs during the cooler parts of the year (possibly favouring vector survival), are positively associated with prevalence. This is compounded by the negative parameter value for temperature annual range (BIO7). A smaller difference between the highest weekly maximum temperature and the lowest weekly maximum temperature had a positive effect on prevalence (presumably through an effect on developmental rates of blackfly immature stages, and perhaps also of the parasite larval stages within the vectors).

The normalized difference vegetation index (NDVI) for the December–February ecological quarter was a statistically significantly and negative predictor of microfilarial prevalence (P-value < 0.05).

This quarter follows a period of low precipitation within the OCP area during which the temperature is approximately half the annual maximum across the northern stretches of the study area, and the highest NDVI values for this quarter are found only along the coastal regions, which are unsuitable as blackfly breeding habitats. When relaxing the BCIs to 75%, NDVI for other ecological quarters became significantly and positively associated with prevalence. Areas with lusher green vegetation during March–May and June–August were positively associated with onchocerciasis prevalence. These greener areas occur within the OCP boundaries, whereas the desert areas situated north of the programme boundaries are extremely arid due to lower rainfall and higher temperatures, with very low NDVI values and negatively associated with the presence of onchocerciasis. Compound topographic index (a proxy for the slope and relative wetness of the land) had a weak effect in the model, as did the precipitation of the wettest quarter (BIO16).

Altitude has been found to be positively correlated with onchocerciasis prevalence in various foci of Cameroon (*229*), Venezuela (*230*), and Guatemala (*231*) up to an altitude of 1,500 metres above sea level, and negatively correlated above that altitude, with most of the proportional increase taking place at altitudes up to 500 m (*229-231*). However, this association may also depend on the type of river and its geomorphological substrate determining vector breeding suitability (and in Venezuela also an association with vector species composition and the relative vector competence of the various simuliid species for *O. volvulus*), and not merely on altitude (*232, 233*). In this study, altitude was not found to be a significant predictor of onchocerciasis. However, this could be partly explained because both rainfall and temperature variables were included in the model, which may remove some of the effect of altitude on prevalence, and the fact that more than 94% of the total land area in the OCP is situated at less than 500 metres elevation.

3.5.3 Model implementation, validation and limitations

In line with other recent approaches to large-scale geostatistical mapping (*131-138, 209, 210, 234, 235*), MCMC methods were implemented to fit a Bayesian statistical model, which uses linear combinations of the survey data (microfilarial prevalence) and a spatial covariance function to obtain predictive posterior estimates at un-sampled locations. The distribution of such estimates readily translates into measures of uncertainty associated with microfilarial prevalence. For those non-surveyed locations, which are distant from those with data, the posterior distribution tends to be more diffuse, indicating greater uncertainty (**Figure 3-12** in **Section 3.3.4** *Map accuracy and predictive probability maps*).

In this study 25km^2 map pixels were used. This was chosen as a compromise between achieving a desirable degree of spatial resolution and minimising computing time, but leads to some loss of short-scale variation in prevalence below the chosen value. This size of individual map pixels means that very short-scale variation between closely situated villages is smoothed over and is likely to be important given the exponentially decaying relationship between distance from the river and biting density/proportion of parous flies that occurs within this distance (*151*).

A consequence of this smoothing effect is that although the predicted surface provides a robust prediction of endemicity at larger scales, it is less capable of representing variations over very short distances, and the estimation of extreme values may be truncated. The output range of prevalence values in those aged 5 years and over in the predictive map was slightly narrower (from 0.5% to 90%) than the input data (from 0% to 95%), a feature of spatial smoothing also reported by other authors (*209, 210, 236*).

The validation statistics compare reasonably well with those presented in other mapping exercises, such as MAP for 2007 (*209*) and 2010 (*210*), which also applied B-MBG procedures but (in the case of MAP 2007) did not find a strong relationship between malaria prevalence (*Plasmodium falciparum* in those aged 2–10 years) and a range of remotely-sensed environmental covariates. (In their update

of a global *P. falciparum* endemicity map for 2010, Gething et al. (*210*) succeeded in identifying environmental covariates as predictors of malaria prevalence.)

Although this study has a lower overall correlation coefficient (0.69 in this case versus 0.82, for the 2007 version of MAP in Africa (*209*)), the sample size of input data was 7 times lower (737 versus 5,307 input data points). The mean prediction error (0.8% vs. 0.8%, indicating a tendency to overestimate prevalence by less than 1%), and the mean absolute error (12% vs. 11%) are comparable. (In the 2010 MAP iteration for Africa, the correlation coefficient between predicted and observed values was 0.86; the mean prediction error was –0.9%, indicating a tendency to underestimate prevalence by less than 1%, and the mean absolute error was 12% (*210*).)

The ability to predict endemicity class correctly was also lower (59% vs. 71% in the 2007 MAP), but there was only a value of 3.5% for serious estimation errors (i.e. non-adjacent class classification) when using four endemicity classes (the hypoendemic class was expanded to include the non-endemic class). The ability of the model to allocate locations correctly into endemicity class partly depends on the boundaries and number of categories chosen to represent endemicity levels (with n endemicity classes there are $(n-1)^2 - n$ chances for misclassification in the $n \times n$ matrix, and hence more endemicity classes are increasingly sensitive to misclassification), and therefore it is not a wholly unbiased metric of map accuracy.

3.5.4 Concluding Remarks

It would be possible to extend the geostatistical model developed here to map the endemicity of onchocerciasis in APOC's countries. APOC has made remarkable progress using REMO to guide the implementation of CDTI (*45*), and kriging maps (based on the spatial dependency component of the model, without inclusion of environmental covariates) have been prepared (Dr J.H.F. Remme, pers. comm.). Application of the methodology described here would make it possible to refine estimation of infection prevalence at locations for which no ground-truth data are available. It would also be

crucial to develop B-MBG tools to jointly map onchocerciasis–loiasis endemicity and co-endemicity (*138, 234, 237, 238*) across West Africa, as the risk of severe adverse effects due to ivermectin treatment in the latter (*239, 240*) presently hinders progress towards the control and elimination of onchocerciasis in the continent by 2020.

Chapter 4. Estimating the number of persons infected and at-risk of onchocerciasis prior to the onset of the Onchocerciasis Control Programme

The overall aims of this chapter are to estimate the number of people living in the OCP borders and surrounding area of West Africa that were at risk and infected with microfilariae of *O.volvulus* prior to the commencement of the control interventions and to calculate a measure of the disease burden caused by onchocerciasis.

4.1 Introduction

In Chapter 3 a map of the spatial distribution of the pre-control prevalence of microfilaria in former OCP areas was presented. This map used a combination of environmental covariates and the spatial autocorrelation present in the available OCP survey data (see **Chapter 2.2** *Pre-control Data Selection)* to produce a map that was better able to estimate prevalence in un-sampled areas. In this chapter this map is used in conjunction with maps of gridded population counts and urban extents, in the countries of the former OCP, estimated to 1975 to produce estimates of the number of persons infected and at risk. The population at risk and infected prior to the inception of control programmes are essential denominators to calculate the initial burden of disease and its changes as a result of such interventions.

4.1.1 Previous estimates of onchocercal disease prevalence

In 1987 and 1995, The World Health Organization (WHO) Expert Committee on Onchocerciasis released technical reports that included estimates of the global numbers at risk of infection, as well as numbers infected and blind through onchocerciasis (*7, 119*). It was stated that 327,000 persons were estimated blind, whilst 17 million were infected. The 1990 Global Burden of Disease (GBD) study, released in 1996 (*241*), estimated that 884,000 disability-adjusted life-years (DALYs) were lost due to onchocerciasis in 1990, whilst the GBD 2000 study estimated that 498,000 DALYs were lost in 2000. The reduction is due in large part to the activities of on-going control programmes, however they are

also based on estimates of the population infected and at-risk that extrapolate from incomplete survey data. As the coverage of surveys has increased, particularly in areas covered by APOC, via Rapid Epidemiological Mapping of Onchocerciasis (REMO) estimates are updated to reflect updated knowledge about the ground truth (*228, 242*).

According to the latest REMO surveys, more than 37 million people would have been infected with O. volvulus (doubling previous estimates) and a further 120 million people live in areas at risk of infection (*242*). With progress in the coverage of rapid epidemiological mapping of onchocerciasis (REMO) surveys, new areas of endemic infection are being uncovered within APOC's intervention area, calling for an updated estimate of the numbers of people suffering from onchocercal infection and disease from the oft-quoted figure of 37 million persons infected. Like-wise, in the OCP areas, because control interventions against onchocerciasis in West Africa started in 1975, a much more limited set of pre-control data were available to inform estimates of infection prevalence, which have never previously been used in a geostatistical model, incorporating environmental covariates to more accurately estimate prevalence in unsampled areas. In addition, great progress has been made in the last 10 years, constructing standardised population denominators, available in global gridded coverage grids, via the Gridded Population of the World (GPW) project and the Global Rural-Urban Mapping Project (GRUMP). Using the map presented in Chapter 3 and the updated, robust population denominators now available updated figures for the number of person infected and at-risk prior to the commencement of control operations are presented.

4.1.2 The Global Burden of Diseases, Injuries, and Risk Factors Study

The Global Burden of Disease (GBD) Study aims to quantify the global and regional effects of disease, injuries and risk factors (collectively termed disability) on population health (*243*). In econometric terms, the study measures the disutility suffered by an individual experiencing a state of

imperfect health, summing across populations. The metric used to measure this disutility is termed the Disability Adjusted Life Year (DALY) and measures the amount of life lost to disease in one year on a continuum from 0 (perfect health) to 1 (death).

The first GBD study was commissioned in 1990 and its preliminary results were published in 1996, providing estimates for 107 diseases and injuries and 10 risk factors (sequelae), in 8 different regions of the world (*243*). Since the publication of the original study there has been an expansion in the availability of newer, richer, more complete data on health, and accompanying development of novel advanced quantitative methods for the analysis of substantive sources of new data. The 2010 Update to the GBD was initiated with the aim of providing a complete systematic reassessment of all available data on disease, injuries and risk factors, particularly utilising new sources of primary data which have become available and more complete since 1990 (for instance in regions where there has been rapid growth in the last 20 years such as India and Asia) (*244*).

For the purposes of the 2010 GBD study, the world has been divided into 21 distinct geographic regions that are as epidemiologically homogeneous as possible. The results of the study provide age- and sex-specific estimates of DALYs for each of these regions for 291 disease and injuries attributed to 1160 disabling sequelae (*245*).

4.1.3 The Disability Adjusted Life Year

The effects of disease are measured using the DALY, which was first introduced by the World Bank in 1993, as a standardised measure of the years of life lost due to disability (*246*). It is defined as the sum of years of life lost (YLL) due to premature mortality and the years of productive life lost through years living with disease (YLD), weighted by disease severity using a disability weight (DW). In its simplest form the DALYs lost to a disease for an individual is:

DALYs = (age- and sex-specific standard life expectancy – age at death) + (duration of disease x disability weight).

Previous GBD studies (GBD 1990 and 2004) used an age weighting that was applied to healthy years of life lived, i.e., years of life lived as a young adult were valued more highly than those lived as a young infant or by someone nearing or exceeding their standard life expectancy. A discount rate was also applied, implying a time-value of preventing disability sooner. GBD2010 recruited a panel of experts, philosophers, economists and others to debate these issues, and it was decided to value years equally at all age groups and to not apply a discount rate in this round of the GBD (*244*).

4.1.4 Disability Weights

Each sequela condition is associated with a disability weight (DW), which lies on an ordinal scale between 0 and 1 representing the amount of healthy life lost. The value of 0, therefore, represents a state of perfect health, whilst 1 would indicate death. Conditions are ranked and assigned a DW by how much they are judged to reduce a healthy person's life experience, so for instance blindness (DW 0.195) is considered more disabling than 'moderate impairment' of distance vision (DW of 0.03). For example, a person with an ideal life expectancy of 75 years, who develops blindness at 60 and lives for 10 years before dying due to the blindness, would have lost 5 years of healthy life through premature mortality and a total of 1.95 years of life lost living with the disability (19.5% reduction in life quality for 10 years), meaning that the individual would lose a total of 6.95 DALYs.

In a major departure from previous GBD studies, which assigned disability weights largely on the basis of the opinion of a limited number of healthcare professionals, the current iteration of the GBD has sought to ameliorate criticism of the DALY metric (*247-251*) by undertaking a complete reassessment of the disability weight system. GBD 2010 carried out household surveys of adults aged 18 and over (*n=13,902*) in face-to-face interviews in Bangladesh, Indonesia, Peru, and Tanzania and telephone interviews in the USA and through the use of an open-access web-based survey (*n=16,328*) to ask paired comparison questions. Respondents were asked to select the person they considered the healthier out of two hypothetical individuals suffering with two different, randomly selected health

states. The final DWs were set using a probit regression analysis of the pooled response data, rescaled to values between 0 and 1 (*252*).

The DW assigned to a health state has a major impact on the number of DALYs lost to a disease caused by that health state. Distance vision blindness is a health state that may be caused by onchocerciasis if left untreated. Previous GBD studies set the DW for blindness at 0.6. The DW derived from the GBD 2010 is 0.195. However, skin disease, which occurs at a much higher prevalence than blindness has increased in weighting from 0.056 to 0.072 and a health state previously only attributable to onchocerciasis, 'Disfigurement: level 2, with itch or pain' has increased from 0.068 to 0.187 (*252*). Unfortunately, the countries and populations sampled in the new GBD2010 studies did not include communities severely affected by onchocerciasis, and the lower DWs ascribed to impaired vision and blindness have been criticized by the Vision Loss Expert Group which was part of the reassessment.

4.2 Methods

To estimate numbers of persons infected and at risk, the maps of pre-control onchocerciasis microfilarial prevalence from Chapter 3 were used as an input numerator. In this chapter I construct a denominator to calculate the number of people infected and define a classification to assign persons in the denominator as being at-risk or not.

To calculate the disease burden, pre-control relationships between microfilarial prevalence and several disease sequaele were modelled. I conducted an exhaustive literature review to try to find as much raw data as possible on pre-control associations between infection and morbidity. At the conclusion of the review, 899 papers were reviewed and I was responsible for reviewing 445 of them. The list of papers I reviewed is attached in the **Appendix E**. Subsequently, the raw data from a multi-country dataset (*253*) became available, which contained detailed individual-level data on onchocercal skin disease, which contained comparable data across multiple countries. The statistical models to

estimate the relationship between infection prevalence and the disabling sequaele were conceived and fitted by Dr. Luc Coffeng and I did not contribute to their development. The dataset used for modelling the link between blindness and microfilarial prevalence was data collected in the Vina Valley in the savanna areas of Cameroon and this model was also conceived and developed by Dr Luc Coffeng. I use these models in order to estimate a disease burden for onchocerciasis and describe them briefly.

4.2.1 Onchocercal skin disease models developed by Dr Luc Coffeng

Generalised linear mixed models fitted by penalised quasi-maximum likelihood (glmmPQL) were fitted in R using data from the multi-country study (*253*). Onchocercal skin disease (OSD) classifications were used as in Murdoch et al (*2*), mapped to the disabling sequelae available in the GBD study, with the worst sequelae assumed to be relevant to a persons' impact on their health.

There is an array of skin lesions resulting from onchocercal infection. These can range from mild, localised, papular rashes to chronic and extensively distributed lichenified, atrophied and depigmented skin. Several classification systems for onchodermatitis exist (*2, 253*). For this study, skin lesions are aggregated into three sequela describing disability suffered from disfigurement, itching or non-specific pain, in three increasing levels of severity, based on (*2*) in order to use with GBD disability weight associated with these sequelae.

Whilst a single person may present with multiple skin lesions from more than one of the severity levels, the sequelae are assumed to be exclusive – i.e. a single individual suffers disability according to the most serious skin lesion they experience. Acute papillary onchodermatitis (APOD) was assigned as the lowest level of OSD (and hence would adjusted by the lowest disability weight score). Individuals displaying depigmentation (DPM) were classified into the second severity level of the GBD (whilst this is a longer-term clinical presentation it is characterised usually by a lack of itching and hence may have less of an effect on a persons' health state). Finally, person suffering chronic

popular onchodermatitis (CPOD) were classified into the third severity level due to a longer association with intense pruritic itching and the co-presentation of other acute lesions caused by onchocerciasis.

All the models were fitted using a random effect for village and household and included covariates for age and sex and interaction terms in the following forms:

$$apod = \beta_0 + \beta_1 * age + \beta_2 * sex + \beta_3 * mf_{5+} + \beta_4 * age * sex + \beta_5 * (age^2) + \beta_6 * (age^2 * mf_{5+}) + \beta_7 * (age * mf_{5+}) + \beta_8 * (age^2 * sex)$$

$$dpm = \beta_0 + \beta_1 * age + \beta_2 * sex + \beta_3 * (age * sex) + \beta_4 * (age^2) + \beta_5 * mf_{5+} + \beta_6 * (sex * mf_{5+})$$

$$cpod = \beta_0 + \beta_1 * \log(\log(age)) + \beta_2 * mf_{5+} + \beta_3 * (\log(\log(age)) * mf_{5+}) + \beta_4 * sex$$

An example of the output of the model for APOD at two different prevalence levels is shown in **Figure 4-1**.

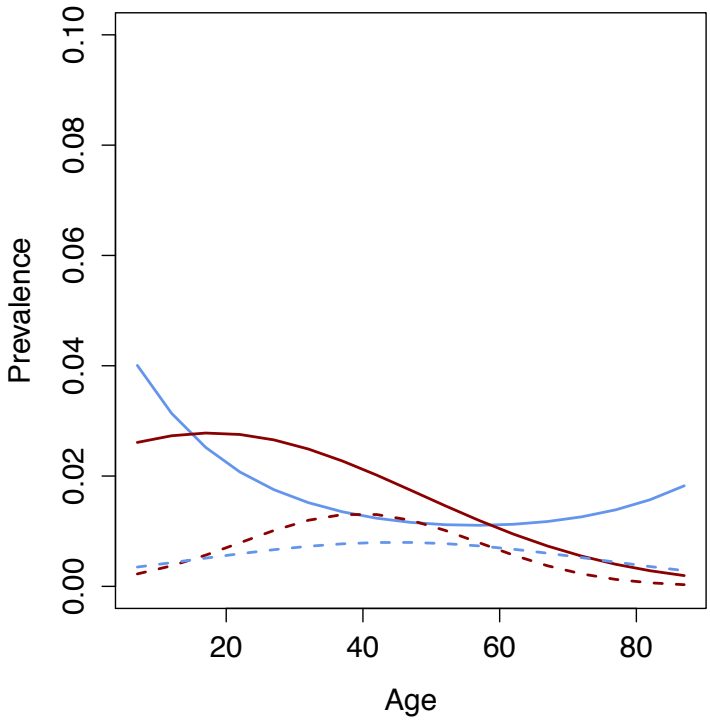


Figure 4-1: Example output from random effect generalised linear model for acute papular onchodermatitis (APOD). Solid lines are age and sex specific output at a prevalence of 90%. Dashed lines are the same model output for an input prevalence of 20%. Blue lines denote male and red lines denote females.

By using the modelled prevalences from Chapter 3 as input to these models, it was possible to calculate an age and sex specific prevalence for each of the three severity levels of onchocercal skin disease. This prevalence was indirectly standardised to a standard population and then multiplied by the overall denominator (total population aged 5 and over in 1975) which is estimated in this chapter, to give sequelae specific estimates of prevalence of morbidity.

These prevalence estimates can be used to directly calculate year living with disease in the GBD 2010 framework. Previously used age-weighting, time discounting and incidence based estimates of the YLD component of DALYs was dropped in favour providing a simpler to calculate, more easily interpreted metric (*244*). Since OSD rarely causes death, in this chapter I shall focus only on the YLD part of the burden of onchocercal skin disease.

4.2.2 Estimating Human Population Counts, Numbers at Risk, and Numbers Infected with *O. volvulus* in 1975

To estimate numbers of persons infected, maps estimating mean endemicity were overlaid on a population count layer to calculate the number of persons expected to have been infected with *O. volvulus* microfilariae. The number of persons at risk was assumed to be the total population aged ≥5 years living within rural areas as onchocerciasis transmission requires clean, unpolluted and fast flowing bodies of water as breeding sites for the vectors. (The selection of the ≥5-year olds was due to the fact that this is the age group examined for skin microfilariae in most surveys and certainly in the OCP epidemiological database. Children younger than 5 years of age are at risk of acquiring onchocerciasis but there was a need to make the population groups comparable to those for which infection with *O. volvulus* is routinely reported.)

Raster global population counts are readily available from the Gridded Population of the World (GPW) v.3 data (http://sedac.ciesin.columbia.edu/data/collection/gpw-v3) from the Center for International Earth Science Information Network (CIESIN) within the Earth Institute at Columbia University (*205*). These data were downloaded for 1990 (the earliest available data) and projected backwards to 1975, on a per-pixel basis, using the relevant national medium-variant, inter-censal country population growth rates (see **Table 4-1**) from the United Nations, World Population Prospects, 2012 Revision (WPP2012) database (*254*). As microfilarial prevalence is often expressed in terms of the population aged 5 years and older, the proportion of the population aged 0–4 years in 1975, according to the WPP2012 database, were removed (see above). The resulting raster was then re-projected and cropped to match the study area covered by the estimated prevalence maps from Chapter 3. **Figure 4-2** depicts the resulting population count.

Total population counts in the estimated 1975 map were very close to the census estimates from the WPP2012 database (**Table 4-2**). The area was delineated into rural and urban areas using the gridded rural-urban extents map (GRUMP v.1) (*206*). Pixels falling in urban extents (**Figure 4-3**) were considered not-at-risk and were excluded from calculations of at-risk and infected totals, as these

areas are not suitable for simuliid breeding. Estimates of prevalence of infection were made within the total study area, and within the OCP control area (all pixels within the white polygonal outline on **Figure 4-2**).

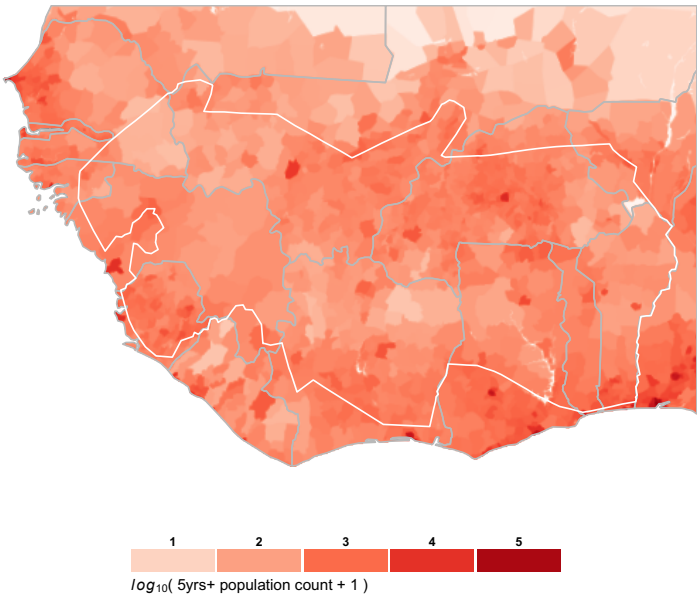


Figure 4-2: Estimated log10+1 transformed population counts for 1975.
Each map pixel contains the log10 transformed count of the total population +1 (to avoid negative values for population counts), aged 5+ estimated to live in that map pixel in 1975. Summing pixels by country, the totals show < 0.8% difference to United Nation Development Programme country census figures for 1975. Grey areas in this map are classified as urban extents.

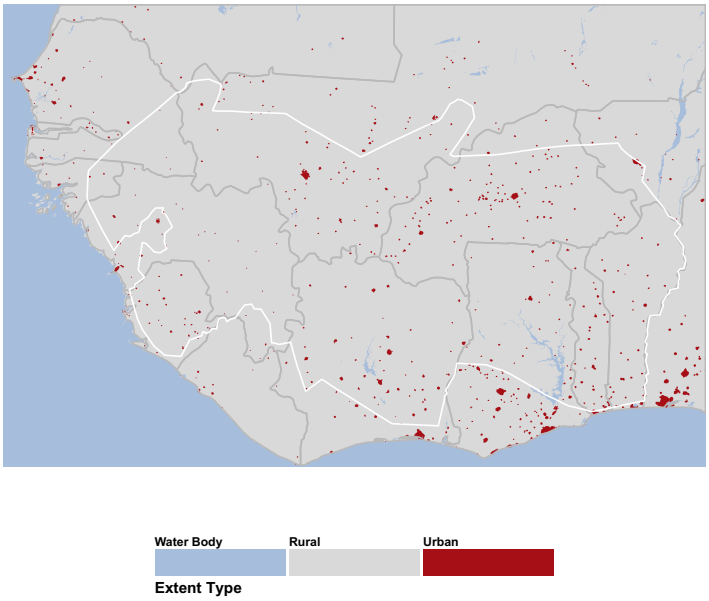


Figure 4-3: Gridded rural-urban extents mask.
Red pixels delineate pixels that are classified as urban extents according the Global Rural-Urban Mapping Project (GRUMP V1.0). Estimates in urban pixels and water bodies do not count towards population totals of numbers infected in Table 4-2.

4.2.3 Backward Projections of Population Counts to 1975

Gridded population counts with a consistently processed global coverage are available from the Gridded Population of the World dataset (GPWv3.0), and the subsequent Gridded Rural Urban Mapping Project (GRUMPv1.0) that build upon GPW and refine and improve it for current population estimates. Unlike in Hay et al. (*209*) where the GRUMP population count data were used, we used the GPW population counts as our reference population from which to back calculate. The argument for using GRUMP over GPW is that firstly the data are available at a finer spatial resolution (1-Km and 5-Km respectively), but as we were not sampling from this dataset, this was not an issue as we were able to use the data at its native resolution for the purposes of estimation. Secondly, the GRUMP data are constructed using a fine-scale areal weighting, based predominantly on night-time lights data collected by the National Oceanic and Atmospheric Administration (NOAA).

The effect of this areal weighting is to better cluster persons into known settlements and out-of sparsely populated, large administrative zones where there is little information to inform the spatial distribution of populations. However, this may under-represent the rural population of West Africa in 1975 using the method of back-calculation applied in this study. The calculations assume a constant rate of rural-urban migration and the effect of using night-light data measured post-1975, where increasing urban populations sizes would make a greater contribution to night-time light signals, would inflate the population count moved to urban areas which in turn would artificially depress the rural population count after applying the back calculation to 1975 population levels.

To back calculate population counts to 1975 totals, a country-specific, 5-yearly-average growth rates was applied on a per-pixel basis. The growth rates were taken from the World Populations Prospects: 2012 Revision [29]. The formula we applied for a given pixel, p, with country-specific average growth rates for the periods 1985-1990 (g_1), 1980-1985 (g_2) and 1975-1980 (g_3) was:

$p_{75} = p_{90} \times \prod_{r=1}^{3} g_r$ with p_{75} and p_{90} being the required population count for a given pixel in 1975, and the known population count for that pixel in 1990 (the earliest observed data). The overall

country-specific population growth rate, g_r, across the complete period 1990-1975, is given by the product of the 5-year average, country-specific growth rates, $\prod_{r=1}^{3} g_r = e^{-g_1 \cdot t} \times e^{-g_2 \cdot t} \times e^{-g_3 \cdot t}$, with $t = 5$. Growth rates are expressed as a percentage (e.g. the 5% growth rate would be 0.05). Because mf prevalence in the OCP is expressed in terms of the population aged 5+, we used The World Population Prospects: 2012 Revision data on Age composition (*254*) in 1975 to remove the proportion of the p_{75} population attributed to this age-class. The country-specific growth rates and proportion of the population aged 0-4 years are given in **Table 4-1**.

Country	g_1	g_2	g_3	$\prod_{r=1}^{3} \exp(-g_r t), t = 5$	Proportion of the population aged 0-4 years
Benin	0.026	0.028	0.031	0.654	0.179
Burkina Faso	0.021	0.025	0.026	0.698	0.179
Ivory Coast	0.045	0.041	0.035	0.546	0.195
Ghana	0.019	0.033	0.028	0.670	0.184
Guinea	0.007	0.024	0.034	0.723	0.166
Guinea Bissau	0.010	0.021	0.022	0.767	0.206
Mali	0.017	0.019	0.015	0.775	0.174
Niger	0.028	0.028	0.029	0.654	0.201
Senegal	0.026	0.029	0.031	0.651	0.19
Sierra Leone	0.024	0.024	0.024	0.698	0.178
Togo	0.024	0.036	0.030	0.638	0.189

Table 4-1: Country-specific medium-variant intercensal growth rates.
Growth rates were obtained from the UNDP World Population Prospects: 2012 Revision (*254*); g_1 refers to the growth rate for 1985–1990, g_2 to that for 1980–1985 and g_3 to that for 1975–1980. The expression $\prod_{r=1}^{3} \exp(-g_r t), t = 5$ is the overall reduction applied to each P_{90} pixel population count in that country in 1990 to obtain the P_{75} pixel population count in 1975.

4.3 Results

4.3.1 Estimated Numbers of People living in the OCP area, Infected with *O. volvulus* and At-Risk of Onchocerciasis in 1975

The population totals calculated in this study are very close to the country-level census numbers provided in (*254*). The total gridded population living in OCP countries in 1975 according to these estimates was 52.5 million, comparing well with the total UN estimate for OCP countries of 52.4 million persons. I combined the map of mean microfilarial prevalence in those aged 5 years and above with a map of the estimated corresponding population for 1975 to derive estimates of numbers of people infected with O. volvulus. I estimated that there were 7.55 million persons with onchocerciasis at the outset of the OCP living within the control programme's boundaries (approximately 1.3 million km2 (*86*)). The total number of persons estimated to be at risk of onchocerciasis within the OCP borders was estimated to be 17.8 million. Within the study area, there were an additional 11.6 million persons aged 5 years old or more, living in rural areas in OCP countries, but outside of the control programme's borders. From the model estimates, I calculated that an additional 2.8 million of these persons may have been living with onchocerciasis in these areas. Of these, 1.1 million persons were in Ghana, where the control programme subsequently extended southwards from the original OCP boundaries to cover forest areas known to be endemic but not previously covered under the OCP. The total numbers of persons at-risk and infected, by country, are shown in **Table 4-2**.

Country	1975 total population count (this study)	1975 Total UNDP population count	At risk population*	Total No. infected‡	Mean prevalence++
Benin	3,061,326	3,262,959	1,591,242	621,600	51.8%
Burkina Faso	6,269,113	6,154,548	3,710,859	1,382,943	40.7%
Ghana+++	10,023,764	9,831,636	2,140,328	1,103,260	52.7%
Guinea	4,386,421	4,350,256	1,676,141	694,441	43.9%
Guinea Bissau	679,563	778,096	106,654	35,042	34.1%
Côte d'Ivoire	6,656,868	6,606,395	2,616,591	1,242,853	52.8%
Mali	6,836,293	6,191,134	3,024,280	1,108,550	37.0%
Niger	5,037,758	5,070,583	227,909	46,299	26.3%
Senegal	4,645,508	4,901,053	122,813	41,388	35.2%
Sierra Leone	2,729,938	2,820,782	1,392,505	708,285	51.3%
Togo	2,149,634	2,410,446	1,161,603	564,466	52.6%
Total	52,476,186	52,377,888	17,770,925	7,549,127	45.2%

Table 4-2: Total population, at-risk population, numbers with onchocerciasis and mean microfilarial prevalence by OCP country in 1975, at the commencement of the program.
* at risk population is defined as the total population aged 5 years and above living in rural areas within the Onchocerciasis Control Program borders. Rural areas are defined by the Urban Extents mask (**Figure 4-3**).
‡ the product of the mean per-pixel microfilarial prevalence and the corresponding pixel at risk-population count. Only pixels within OCP borders count towards this total.
++ the average mean prevalence values by country, in percent. Country mean prevalence is the average of the per-pixel mean prevalence from the map output depicted in Figure 3A, within OCP borders (not the ratio between the numbers of infected and the numbers at risk).
+++ the values shown here do not include a large area in southern Ghana later incorporated into the OCP and that mainly comprised forest-type onchocerciasis.

4.3.2 Estimating disease burden caused by onchocerciasis

Using the models developed by Dr Luc Coffeng in conjunction with the output of my microfilarial prevalence maps it was possible to make estimates of the burden of skin disease in West Africa. Depicted in **Table 4-3** is the total number of cases of the different types of skin lesions shown by country for the pre-control population that were resident in 1975. Ghana, the Ivory Coast and Nigeria (referring to just the portion of Nigeria in the study area) had the highest number of cases of skin disease.

Country	APOD (Females)	APOD (Males)	CPOD (Females)	CPOD (Males)	DPM (Females)	DPM (Males)	Total
Senegal	6,574	5,076	2,523	6,198	15,540	19,051	54,962
Guinea Bissau	1,182	937	558	1,355	2,940	3,441	10,413
Gambia	810	634	346	844	1,974	2,358	6,966
Mauritania	1,425	1,069	417	1,044	3,156	4,093	11,205
Guinea	11,414	9,556	8,694	20,229	30,036	32,599	112,528
Sierra Leone	8,123	7,021	7,545	17,404	22,124	22,938	85,155
Liberia	5,053	4,212	3,505	8,304	13,492	14,609	49,174
Mali	15,836	12,937	9,812	23,303	40,762	45,691	148,342
Ivory Coast	16,419	14,101	14,657	33,848	44,482	46,525	170,032
Burkina Faso	15,793	13,259	12,273	28,530	41,658	45,045	156,558
Ghana	22,297	18,891	18,066	42,031	59,794	63,652	224,731
Togo	5,663	4,947	5,714	13,062	15,472	15,854	60,711
Niger	3,295	2,559	1,330	3,257	7,876	9,557	27,875
Benin	7,038	5,995	6,133	14,150	18,667	19,839	71,822
Nigeria	17,461	14,801	13,984	32,689	47,021	49,927	175,883
Total	138,382	115,995	105,559	246,248	364,994	395,179	1,366,357

Table 4-3: Total number of cases of onchocercal skin disease by country in 1975.
Cases falling in urban areas have been removed from this total.

I used the disability weights from the GBD 2010 study to estimate the number of years of healthy life lost through living with disease. For skin disease severity 1 the disability weight was 0.029, for skin disease severity 2 the disability weight was 0.187 and for skin disease severity level 3 the disability weight was 0.562. The total number of YLDs lost in 1975 due to OSD was 347,244 (**Table 4-4**). An example of the age specific output, for a selection of ages, from the each of the three models of skin disease is shown in **Figure 1-4**, **Figure 1-5** and **Figure 1-6**.

Country	APOD (Females)	APOD (Males)	CPOD (Females)	CPOD (Males)	DPM (Females)	DPM (Males)	Total
Senegal	191	147	1,418	3,483	2,906	3,563	11,707
Guinea Bissau	34	27	314	762	550	643	2,330
Gambia	23	18	194	475	369	441	1,521
Mauritania	41	31	234	587	590	765	2,249
Guinea	331	277	4,886	11,369	5,617	6,096	28,576
Sierra Leone	236	204	4,241	9,781	4,137	4,289	22,887
Liberia	147	122	1,970	4,667	2,523	2,732	12,160
Mali	459	375	5,515	13,096	7,623	8,544	35,612
Ivory Coast	476	409	8,237	19,023	8,318	8,700	45,163
Burkina Faso	458	385	6,898	16,034	7,790	8,423	39,987
Ghana	647	548	10,153	23,621	11,182	11,903	58,053
Togo	164	143	3,211	7,341	2,893	2,965	16,718
Niger	96	74	748	1,831	1,473	1,787	6,008
Benin	204	174	3,447	7,953	3,491	3,710	18,978
Nigeria	506	429	7,859	18,371	8,793	9,336	45,295
Total	4,013	3,364	59,324	138,391	68,254	73,898	347,244

Table 4-4: Total number of healthy years of life lost due to years living with disease of onchocercal skin disease (OSD) in the rural popualtion aged 5+ in 1975.

Ghana had the highest burden of OSD when measured by the metric of years living with disease. The biggest loss of healthy life occurred in males who were suffering from chronic popular onchodermatitis.

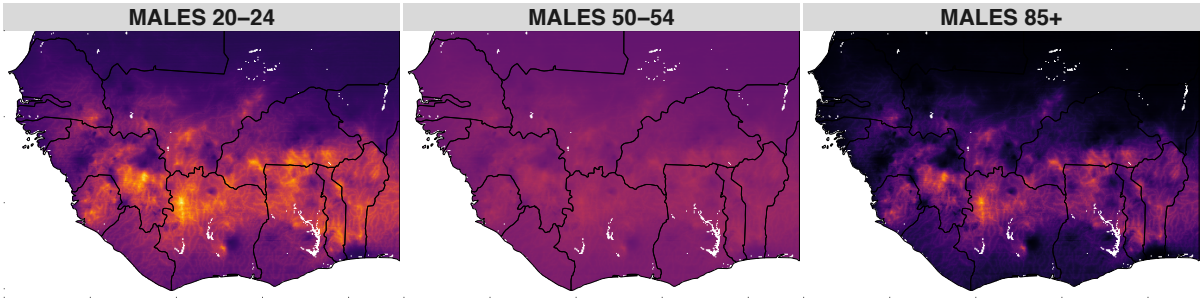


Figure 4-4: Age specific prevalence of APOD.
Calculated from the acute papular onchodermatitis (APOD) model for three different age ranges. APOD reduces in prevalence over time as infected individuals move to more severe OSD sequelae.

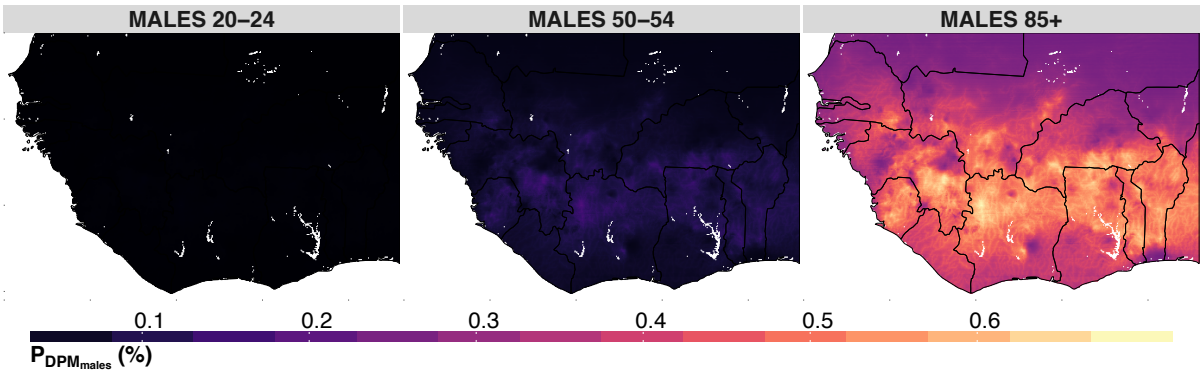


Figure 4-5: Age specific prevalence of DPM.
DPM increases in prevalence over time as this morbidity develops as a result of prolonged and cumulative exposure to infective bites from blackflies.

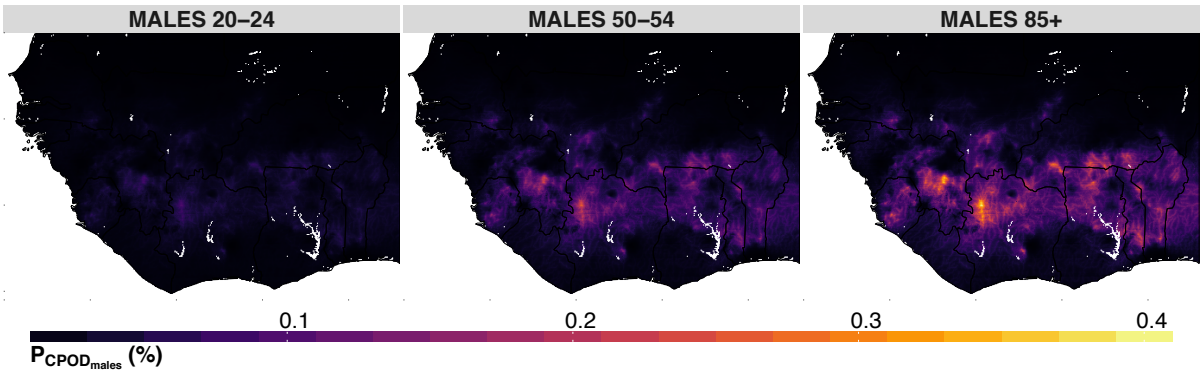


Figure 4-6: Age specific prevalence calculated from the chronic papular onchodermatitis (CPOD) model.
This is the most severe form of skin disease, characterised by both itching and lesions. It is however transient as people move to the less severe DPM sequelae.

4.4 Discussion

This study is, to the best of my knowledge, the first estimate of the numbers of persons infected with - and at risk of- onchocerciasis, using robustly estimated numerators of infection prevalence from a model-based geostatistics study, and using per-pixel estimated denominators of at-risk populations in these areas prior to the start of control interventions.

4.4.1 Population at risk and numbers infected with *O. Volvulus* in 1975

The United Nations estimated the 1975 population for the 11 countries covered by the OCP at 52.4 million (*254*). The back-calculation used in this thesis estimates this figure at 52.8 million (only 0.8% higher). Boatin (*86*) states that by the time the OCP ended in 2002, 27 years later, the estimated population in the area covered by the programme was over 78 million.

The Third Report of the Expert Committee on Onchocerciasis, published in 1987 (*119*), stated that in the OCP area, the number of persons at risk (as assessed in 1984–1985) was 7.2 million, with 2.4 million infected (a ratio of 3:1). Since this assessment was made approximately 10 years after the commencement of anti-vectorial operations in the seven initial countries of the OCP (namely, Benin, Burkina Faso, Côte d'Ivoire, Ghana, Mali, Niger and Togo), the report excluded persons living in previously endemic areas of these countries where transmission was deemed to be controlled from the calculations of the numbers at risk. In those areas, the numbers of persons infected with *O. volvulus* were calculated as 25% of those who would have been infected at the beginning of the control programme (*119*).

Consequently, the numbers infected at the start of the OCP would have been 5.9 million in the initial 7 countries, and 6.8 million adding Guinea, Guinea Bissau, Senegal, and Sierra Leone (which were incorporated later). According to this mapping study, the numbers of persons at risk would have been 2.5 times as large as the number reported in 1987, at 17.9 million (representing the total population

aged 5 years and above living in rural areas within the OCP borders). The number of people with onchocerciasis would have been 7.6 million, i.e. three times as large as the number documented in 1987 (which gives directional support to the report's consideration that the figures given would have been 25% of those prior to control).

The report also acknowledges that the figures for the prevalence of infection most likely underestimated the true extent of the problem, as they were obtained by extrapolation from relatively small sample surveys, without accurate knowledge of the local distribution of endemic foci (and certainly without considering spatial correlation between survey points).

Survey villages were selected from historical records, for their hyperendemicity of onchocerciasis, and other criteria aimed to ensure long-term evaluation of vector control. Village selection was somewhat independent of the entomological evaluation network, and consequently, although successful in evaluating the vector control campaign, it could not have described the distribution and severity of the disease adequately and did not allow for correct estimates to be made of the number of people infected with *O. volvulus* (*121*).

A study of the distribution of onchocerciasis was conducted in 1987 in the Western Extension (*121*), partly conceived to overcome these limitations. After discarding population sizes greater than 800 inhabitants, the authors calculated a rural population of 4.5 million persons and reported that, of these, 1.5 million persons were infected. Using the Gridded Population of the World dataset, and satellite-derived GRUMP data as a clipping mask to remove urban extents, the model used in this thesis estimated that the rural population at risk within the Western Extension in 1975 would have been 11.4 million, with 3.9 million persons infected. Differences may partly be due to the more stringent exclusion criterion used in (*121*). Although absolute numbers differ, the ratio of at-risk persons to infected persons is the same (3:1) in (*121*) and this thesis.

Finally, Coffeng et al. (*125*) estimated that before the inception of APOC in 1995 about 32 million people in APOC areas were infected with onchocerciasis. Adding this figure to the 7.6 million infected prior to the inception of the OCP estimated from this study, a total value of nearly 40 million

people infected in Africa at baseline is obtained. This total lends support to the recent estimation of 37 million people infected with *O. volvulus* in the continent (*10, 255*), and supersedes the frequently quoted but grossly underestimated previous figure of about 18 million infected (*119*).

4.4.2 Burden of disease due to onchocercal skin disease

Comparing the results from this study to the latest available on the Institute for Health Metrics and Evaluation's Global Health Data Exchange (GHDx) website (http://ghdx.healthdata.org/gbd-data-tool: accessed 7th June 2016) there is good concordance between the two. For instance, in Ghana, in 1990, the latest GBD update estimated that there were 56,148 years of healthy life lost in Ghana for 1990. I estimated in this study, that in 1975 Ghana suffers the loss of 58,053 healthy years due to years living with onchocercal skin disease.

A major difference between this study and the estimates of the GBD however, is that the GBD study only consider pruritic itch as an all-encompassing disease sequelae at disease severity level 2. It could be argued that this classification does not capture the full range of disability that can be caused by onchocercal skin disease. The skin lesions arising from onchocercal infection can range from mild, localised, papular rashes to chronic and extensively distributed lichenified, atrophied and depigmented skin. The skin can often covered in sores and pustules as a result of intense pruritic itching, which in itself can be chronic and lead to psychological and socio-economic problems, including social isolation, insomnia and absenteeism from work (*55-58*). A single disease sequelae may not be able to adequately capture the full range of disability which is a major drawback of how onchocercal skin disease is currently measured in the GBD framework.

4.4.3 Limitations

One of the major limitations of this study is that I considered only years living with disease, rather than the additional burden of years of life lost. However, the burden of years of life lost due to OSD is not thought to be significant. An obvious extension to the work undertaken here would be to incorporate an estimate of the YLDs and YLLs due to low vision and blindness, which would be substantially greater than the 347,244 YLDs estimated here. Another limitation and extension to this study would be to include a correction factor for bioclime, so that the relationship between forest type OSD symptoms and savanna OSD symptoms can be estimated separately.

Chapter 5. Socioeconomic analysis of factors associated with onchocerciasis and compliance to community-directed treatment with ivermectin.

5.1 Chapter Aims and Objectives

The overall aim of this chapter is to identify factors associated with infection with onchocerciasis and factors associated with compliance to community directed treatment with ivermectin (CDTI) in Ghana. Compliance is defined as short-term compliance (compliance with CDTI at the previous treatment round), and compliance type (either good compliance, random-non-compliance or systematic non-compliance).

5.2 Introduction

Ivermectin is a highly effective microfilaricide which is able to eliminate high numbers of microfilariae in the skin of infected hosts within 48 hours of ingestion, with almost complete clearance found 4 weeks after treatment (*256*). Subsequently, repopulation of skin microfilariae is usually not detected for a period of 6-9 months (*257-259*). At the standard dose of 150μg/kg bodyweight ivermectin does not possess a potent macrofilaricidal effect against adult *O. volvulus,* however it substantially reduces the fecundity of adult female worms for a period (*260, 261*), with a cumulative and permanent reduction in fecundity contributing to the low rates of repopulation of skin microfilariae. This 'embryostatic' effect means that ivermectin is able to interrupt disease transmission given sufficient coverage in a community.

Theoretically for the goal of elimination, the frequency of drug distribution (preventing inter-treatment transmission) and duration of distribution (longer than the natural lifespan of adult worms) must be sufficient to eliminate all life stages of the parasite from host and vector populations. In Ghana, despite apparently high levels of coverage over a prolonged period, there are many communities where infection persists, transmission is on-going (*262*), and suboptimal responses to ivermectin are being observed (*110*). Some of these communities were the first to receive ivermectin in 1987, and have received more than 25 annual doses of ivermectin. A longitudinal epidemiological

study conducted by Osei-Atweneboana *et al*. (*113*) in these communities in Ghana, found some villages with a long treatment history now respond sub-optimally to treatment with ivermectin.

The rate of mf repopulation in the skin is linked to the infection intensity and therefore the parasite establishment rate (PER). It is positively correlated with the initial endemicity and intensity of infection (*184*). It is negatively correlated with the total number of treatments received by an individual due to the cumulative and permanent reduction in fecundity of adult females by ivermectin (*106*), although this embryostatic action of ivermectin has been disputed in subsequent studies (*104, 263*).

In the Osei-Atwenenboa follow-up study(*113*), the rate of repopulation of skin microfilariae in sub-optimally responding villages could be as much as 50% of pre-treatment loads after 6 months, even though they had received some of the highest numbers of treatment rounds. The rate of mf repopulation in normally responding villages is expected to be around 20% 1 year after treatment (*103, 112*). This 'suboptimal response' to treatment with ivermectin is of serious concern as ivermectin is currently the only drug in widespread use to control onchocerciasis.

In a recent study on measuring the effect of increasing the treatment frequency on suboptimal responses to ivermectin, there was no apparent association between repopulation rates, number of rounds of treatment received, coverage or prior infection intensity (*264*). A reduction in compliance could therefore be responsible for higher than expected rates of repopulation of microfilariae. Identifying individual-level factors that affect coverage and compliance, and thus prolong the required duration of treatment is crucial to the successful outcome of treatment programmes and preventing the recrudescence of disease or emergence of parasitological resistance.

Whilst there have been previous studies that have investigated factors linked with coverage and compliance with ivermectin in Africa, these have been based in APOC countries (*265-267*) and as such may not be applicable to the West African setting of Ghana which was part of the original Onchocerciasis Control Programme (OCP) and has been undergoing vector control since 1975 and ivermectin distribution since 1987.

Figure 5-1: Daily activities of a blind man from Asubende.
This man (who gave permission for his photo to be taken) made a living mending fishing nets and worked a small farm nearby. He was reliant on local help to navigate to the farm, but once there, could work alone.

5.3 Methods

To investigate factors affecting coverage and compliance with ivermectin I planned and carried out an extensive household survey in the Brong-Ahafo and Northern regions of Ghana. The survey was conducted in partnership with the Institute of Statistical, Social and Economic Research (ISSER) at The University of Ghana (UoG) in Accra, Ghana. The survey instrument was a hierarchical, cluster-design random sample survey which was administered to 280 households in 13 communities, covering 1383 household members. A previous version of the survey instrument was previously used by ISSER to collect data on economic change in Ghana. The questionnaire for this study was specially modified to collect information on access to onchocerciasis treatment, whether ivermectin is taken regularly when scheduled in the community, and whether the treatment is working when taken.

5.3.1 Community selection and description of study area

The study was conducted in 13 onchocerciasis endemic communities in savannah regions of Ghana, located in the Brong-Ahafo and Northern regions of Ghana. Most villages were selected from those included in a previous study by Osei-Atweneboana *et al*. (*113*) which had indicated that in some

villages there was cause for concern regarding response to community-directed treatment with ivermectin. There were 7 villages which had microfilarial repopulation rates indicating normal responses to treatment with ivermectin and three villages which had sub-optimal responses to treatment with ivermectin - faster than expected rates of repopulation of microfilariae in the skin of infected hosts. The communities had received (at the time) between 6 and 18 rounds of annual ivermectin treatment (*113*).

In addition, using GIS and parasitological survey data I identified several potential 'transmission hotspots' in Ghana where ongoing transmission was a problem, but where no data on microfilarial repopulation rates were available to positively identify suboptimal responses to ivermectin. I mapped and analysed parasitological data collected from the end-of-programme surveys carried out by the Onchocerciasis Control Programme (OCP) in Ghana during the period 2000-2002. Maps of community microfilarial load (CMFL) and skin snip prevalence from these OCP surveys were compared to maps of nodule prevalence from the latest available (at the time) REMO surveys conducted by the Ghana Health Service (GHS) Neglected Tropical Diseases Control Program (NTDCP), to find villages where transmission appears to be persisting (See **Figure 1-2**).

The poorly responding communities identified in the Osei-Atweneboana *et al*. study were Kyingakrom (Chingakrom) and New Longoro located in the Kintampo District of the Brong-Ahafo region, and Jagbenbendo (Jagbingbindo) in the Kpandai District (formerly part of the East Gonja District) in the Northern region. Also of concern in the Kpandai district was the village of Wiae. However, in a subsequent study the repopulation rates of microfilariae in this village was not found to be statistically significantly different from repopulation rates in a treatment-naive control dataset (*110*).

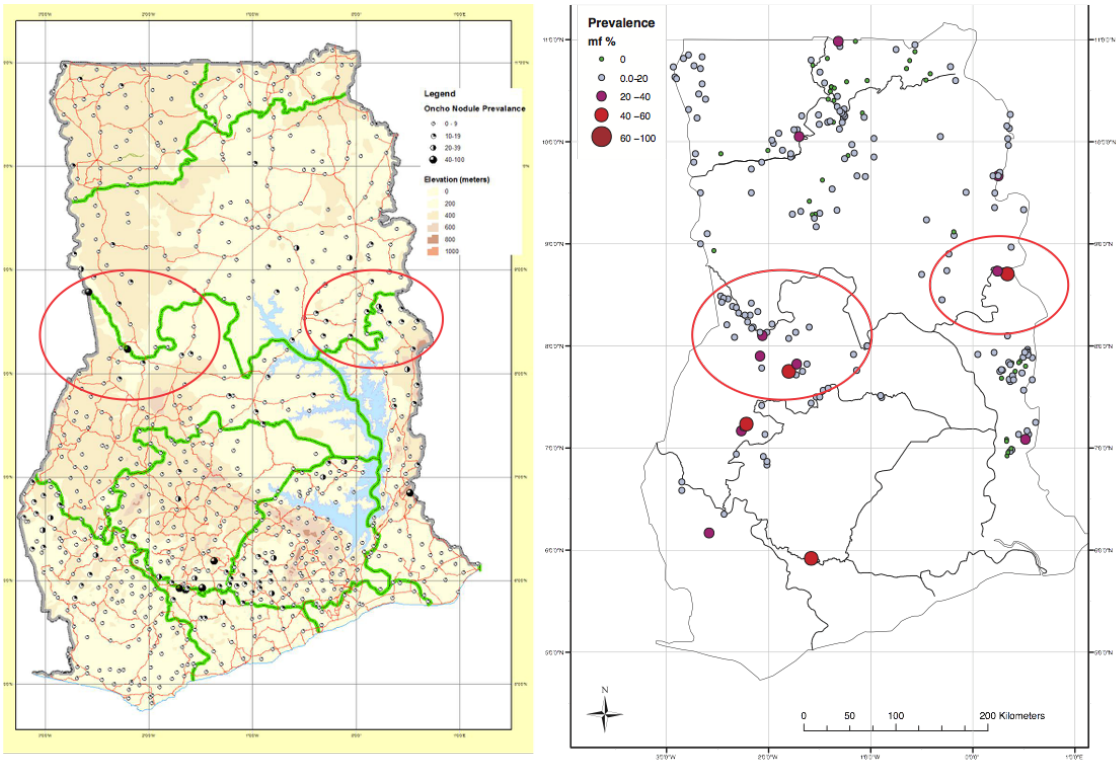


Figure 5-2: Comparison of REMO maps from Ghana Health Service surveys, with end of programme OCP survey data. The clusters outlined in red on either map indicate areas where transmission appears to be ongoing.

From the analysis of OCP parasitological data and GHS REMO surveys I also selected the communities of Agblekeme, Cache and Nsunua in the Bole district of the Northern region where there was persistent infection prevalence. The normally responding communities in Brong-Ahafo were Asubende (an OCP sentinel site that has received ivermectin since 1987), Senyase, Baaya and Ohiampe, all located in the Pru district. I also selected Ntereso from the Bole district, and Takumdo from Kpandai district in the Northern region as normally responding villages on the advice of Mike Osei-Atweneboana (*pers. comms.)*. A map of the selected villages is depicted in **Figure 5-3**.

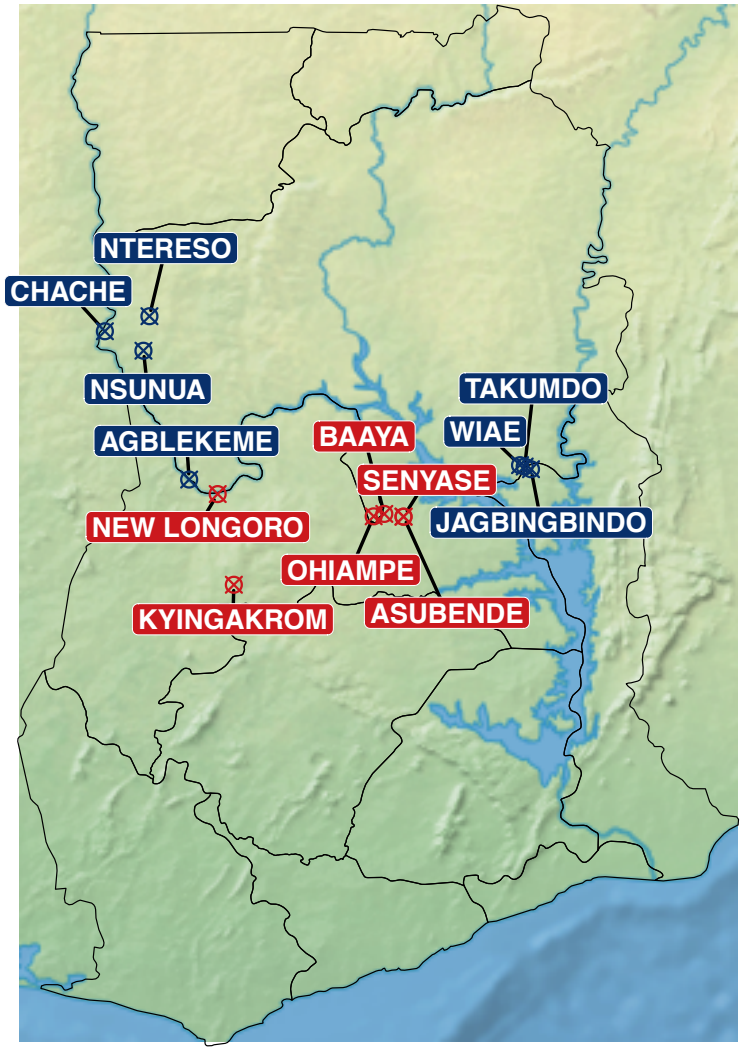


Figure 5-3: Map of selected village locations.
Labels are coloured according to Region each village falls in. Red is Brong Ahafo and Blue is the Northern Region.

5.3.2 The survey instrument and data collection

The household survey was used to collect detailed data on individual household members. The survey was divided into sections comprising demographic history, migration, health in the last four weeks, health in the last two months, health related to onchocerciasis, education history and training and employment history and community participation. The survey was also used to collect data on household assets, including possessions and livestock, and information about the environment the household lives in and the services they have access to.

5.3.3 Household selection and survey administration

Prior to the commencement of the fieldwork a number of skilled interviewers were trained in the use of this survey, and in particular the collection of data pertaining to onchocerciasis infection and treatment. The interviewers each received 2 days training at ISSER in the administration of this study. Each selected community was then visited and an audience sought with the Chief of the village. Once permission was granted for the study to be carried out a listing of all the households in the community was taken.

A roster of the households in each community was created by walking around each community and listing all the homes in the community, using their postal house number. Households were then picked at random from the list and permission was sought from the household head to hold an interview the following day. If the head of the household was not available, the spouse or next of kin's permission was sought. If they were not present or they indicated they were unable to be present the following day a new household was selected at random from the roster. In each community, a total of 20 households were randomly selected, except in New Longoro for which, being a much larger community, 40 households were selected. The interviews were carried out by a total of 5 interviewers who each conducted 4 surveys in each village.

Figure 5-4: Photographs of the interview team at work on this study.
Clockwise from top-right; training of the interviewers at ISSER; Conducting an interview in Ohiampe; Conducting an interview in Senyase; Nightly cleaning and editing of the data helped to reduce errors and omissions occurring at interview. The data were checked for internal consistency.

5.3.4 Data entry and analysis

All original paper copies of the survey instrument were checked at the end of each surveying day by the candidate and the interviewers and any issues arising or inconsistencies were highlighted and corrected as necessary (e.g. conflicting marital statuses between household members). I then entered the edited and cleaned original survey data into a specialised survey data entry platform. I created the data entry application using CSPro, a program of the US Census Survey Bureau.

Initially I created a data dictionary, describing every question on the survey and the possible answer sets that could be given for a given question on the survey instrument. I coded a set of logic rules for many of the questions to prevent incompatible answers from being keyed in during data entry (automatic key-in verification). An example is entering someone as married on the household roster

will prevent the data operator from selecting the household member as ‘widowed’ on a later question. The logic rules also define an automatic skip pattern whereby certain sections of the questionnaire are skipped depending on previous responses to questions, e.g. when a household member is less than 7 they are not asked questions about their work status. Finally, a set of data entry forms were created to enter the data on the computer. See **Figure 5-5** for an example of the data dictionary, the logic coding and the data entry forms.

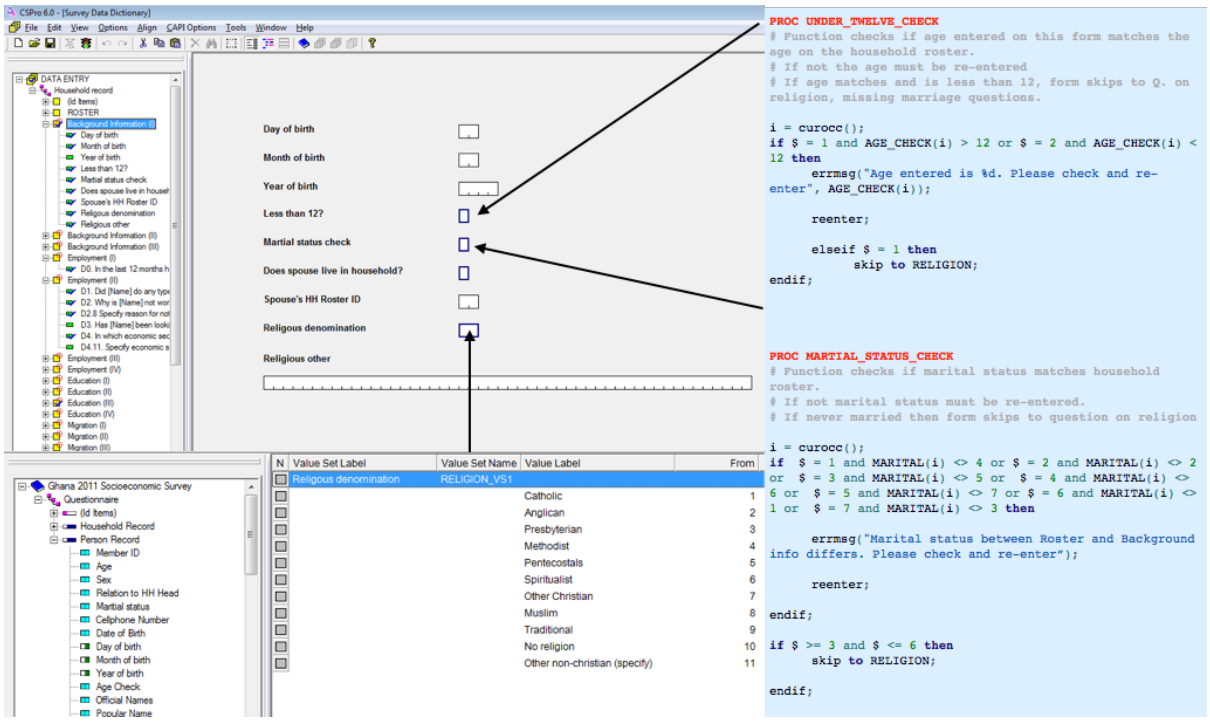


Figure 5-5: Example of the CSPro data entry program.
I used CSPro (US Census Bereau) to enter the raw data. A data dictionary specifying allowable values for every question was constructed. Elements from the data dictionary were used to build data entry forms. Certain elements of the data entry forms were augmented with logic code which prevented the entry of certain answers based on prior answers and defined a skip pattern when conditions were met (e.g. in example above entering a respondents age as ‘12’ skips questions on marital status).

5.3.5 Ethical approval

Ethical approval was obtained from the ethics review committees of the Noguchi Memorial Institute for Medical Research, Ghana [NMIMR-IRB CPN 032110-11], the Ghana Health Service [GHS ERC 04_3_11], and Imperial College London Research and Ethics Committee [ICREC_11_2_4]. Informed verbal consent was obtained from every household after explaining the purpose of the study. Privacy

and confidentiality of the responses were ensured. The original copies of the household surveys were transported back to Imperial College London for long-term storage.

5.4 Results

Coverage data were obtained from the Ghana Health Service to ascertain treatment coverage in the selected communities in the years prior to the study. See **Table 5-1** for details of each village and the recorded treatment coverage.

Region	District	Village	2004	2005	2006	2007	2008	2009	2010 (1)	2010 (2)	2011 (1)	2011 (2)
BRONG-AHAFO	KINTAMPO NORTH	NEW LONGORO					0.581	0.631	0.895		0.892	0.905
BRONG-AHAFO	KINTAMPO SOUTH	KYINGA-KROM						0.786	0.869		0.966	0.847
BRONG-AHAFO	PRU	OHIAMPE			0.738	0.765	0.835	0.839	0.795	0.875	0.782	
BRONG-AHAFO	PRU	ASUBENDE			0.746	0.746	0.875	0.857	0.911	0.903	0.875	
BRONG-AHAFO	PRU	BAAYA			0.691	0.785	0.815	0.875	0.842	0.893	0.807	
BRONG-AHAFO	PRU	SENYASE			0.746	0.746	0.782	0.799	0.814	0.846	0.822	
NORTHERN	BOLE	NTERESO	0.826	0.697	0.883					0.937		
NORTHERN	BOLE	AGBLE-KEME	0.868	0.775	0.713				0.948	0.948		
NORTHERN	BOLE	CHACHE	0.928	0.761	0.822				0.894			
NORTHERN	BOLE	NSUNUA	0.813	0.722	0.852				0.820			
NORTHERN	KPANDAI*	TAKUMDO	0.809	0.782	0.702				0.820		0.858	0.754
NORTHERN	KPANDAI*	JAGBING-BINDO	0.957	0.879	0.927				0.820		0.842	0.777
NORTHERN	KPANDAI*	WIAE		0.930	0.785				0.900		0.944	0.806

Table 5-1: Overall treatment coverage for 7 years prior to this study.
Missing data indicate that the data were not available or that a treatment round did not take place in that location.

5.4.1 Overview of the household survey

I entered data for 1383 individuals in 280 households from the 13 villages. Of these, 1042 complied with ivermectin treatment within the last 12 months, whilst 336 did not, and 5 were of unknown status with respect to compliance. Of the total study population, 1151 persons (83.2%) were eligible for treatment, meaning 90.53% of eligible persons took ivermectin, and the crude treatment coverage across the total study population was 75.3% (including those who were ineligible). The median household size per village ranged from 3 persons in Kyingakrom to 6 persons in Wiae with an overall study median household size of 5. There was no difference in average household size between the Northern and Brong-Ahafo regions. The main respondent (nominally the head of the household) for each survey was asked 'Are you happy with your life?', to which 237 (84.6%) responded 'Yes', 41 responded 'No' and 2 households did not answer. There was no association between how heads of households responded to this question and their self-reported infection status with onchocerciasis (χ^2 0.592, OR 1.34, 95% CI 0.64-2.81, p-value 0.442).

5.4.2 Age, sex and family status

The 1383 respondents were 49.8% female. The mean age of females in the study sample was 20 years 7 months, and males had a mean age of 21 years 9 months (median ages of 15 and 16 respectively). The reported age range was from under 1 years old to 99 years (90 years in females). A histogram of the reported ages revealed spikes at 5 yearly intervals (e.g. 20, 25, 30 etc.) indicating possible recall bias or rounding in reported ages.

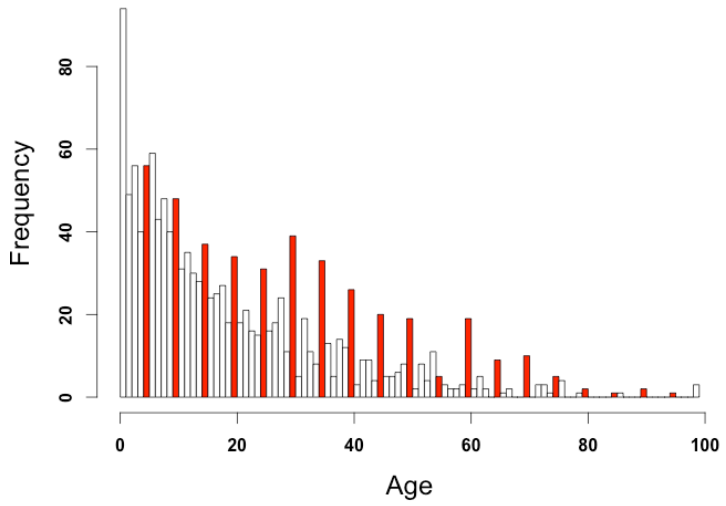


Figure 5-6: Age distribution of study population.
Each bar represents 1 year. Every 5th year (e.g. 20,25,30) is coloured red, showing over-reporting of these ages, compared to nearest other ages.

Of the 818 persons aged 12 years and above, 316 (38.6%) reported as having never been married and by age 25 this fell to just 41 of 490 (8.37%) persons reported as having never married. At the time of the study 398 persons were currently married (28.8%). There were 53 babies aged under 1 year of age, giving a crude birth rate (excluding mortality of infants under 1 year old in the past 12 months) of 3.83% (or approximately 38 live births per 1000 population). The full breakdown of marital status across two regions is given in **Table 5-2** below.

Marital status	Northern	Brong Ahafo	(all)
Married	227	169	396
Never married	159	161	320
Consensual union	15	30	45
Divorced	6	9	15
Separated	0	3	3
Widowed	12	27	39
(all)	**419**	**399**	**818**

Table 5-2: Marital status of all study participants aged 12 years and older.

The relationship to the designated head of household for all respondents (e.g. household head, spouse, child, foster-child etc.) is shown in **Table 5-3**. In one case an interview was carried out where no-one

identified as the household head, giving 279 heads of household, rather than 280. The average number of children per household was 2.8.

Relationship to head of household	Northern	Brong Ahafo	(all)
Household Head	139	140	279
Spouse	119	88	207
Child	414	366	780
Grandchild	11	30	41
Parent/Parent-in-law	0	4	4
Son/Daughter-in-law	0	2	2
Other relative	7	34	41
Adopted/Foster/Stepchild	10	6	16
Unknown	13	0	13
(all)	**713**	**670**	**1,383**

Table 5-3: Household status of household members from 280 households across the two regions in the study.

5.4.3 Ethnicity, religion and migration

1375 respondents (99.4%) were Ghanaian by birth with 1369 persons (98.99%) born in Ghana. The other respondents were naturalised Ghanaian (3), Togolese (2), Ivorian (1) and Malian (1). The 5 largest ethnic groups in the study were: Kokomba (551 [39.8%]); Dagaaba (181 [13.09%]); Mo (143 [10.34%]); Ewe (121 [8.75%]) and Gonja (93 [6.72%]). People of Kokomba descent were found equally in the Northern and Brong-Ahafo regions (19% and 21% respectively).

The majority of persons reported as being of Christian faith (56.3%), with the largest denomination being Catholic (22.4%). The next largest religious groups were Traditional (18.1%) and then Muslim religion (13.5%). Traditional religion was more observed in the Northern region where 24.8% of people from that region reported being of Traditional religion, versus 10.7% of people from the Brong-Ahafo. In the Brong-Ahafo region 21.7% of people identified as being of a Christian denomination other than Catholic, Anglican, Presbyterian, Methodist or Pentecostals versus just 4.4% in the North, where Christians were mainly Catholic and Pentecostal. Observance of Muslim religion

was also greater in the Northern region (17.3%) than in Brong-Ahafo (9.3%). There were 11.2% of people reported as not observing any religion at all.

	Northern		Brong Ahafo		(all)
Religion	**#**	**Regional proportion**	**#**	**Regional proportion**	**#**
Catholic	173	0.24	132	0.20	305
Presbyterian	1	0.00	15	0.02	16
Methodist	0	0.00	24	0.04	24
Pentecostals	108	0.15	140	0.21	248
Spiritualist	3	0.00	2	0.00	5
Other Christian	31	0.04	142	0.21	173
Muslim	122	0.17	61	0.09	183
Traditional	175	0.25	70	0.10	245
No religion	91	0.13	61	0.09	152
Other non-christian (specify)	1	0.00	7	0.01	8
NA	8	0.01	16	0.02	24
Christian (any)	313	0.44	453	0.68	766
Non-Christian	301	0.42	140	0.21	441
(all)	**713**		**670**		**1,383**

Table 5-4: Number and proportion of people within each study region, who observe a religion.

95.5% of persons did not spend any time away from their communities in the last year. Of the 62 (4.5%) persons who did spend time away, only 22 were away for a significant period (greater than 3 months) and 10 were away to attend school. Information on migration was collected for 1028 persons aged 7 and over. Of these, 663 (64.49%) were born in the community in which they currently lived. For the 365 persons aged over 7 who were not born in their current community, the median time since migrating there was 12 years, and the mean time since arrival was 15 year and 6 months.

5.4.4 Employment and household possessions

Of the sample, 676 were of working age (defined as between the ages of 15 and 64) with an employment rate of 78.25% (529). The reported average annual pay of those in employment was 962 (Ghanaian Cedis) GH₵ with significantly different levels of pay (Wilcoxon rank sum, W=55345, p-value=<0.0001, shift=160, 95% CI=60-250) between the Northern region (median = GH₵ 500) and the Brong-Ahafo region (median = GH₵ 200). The majority of those working were employed or worked in agriculture, equally across the two regions. A table showing median income and number of people employed by economic sector and region is shown in **Table 5-5**. The distributions of income for those working was also compared for those reporting suffering from onchocerciasis versus non-sufferers. Employed onchocerciasis sufferers appeared slightly better off with respect to income than those not reporting onchocerciasis (Wilcoxon rank sum, W=37494, p-value = 0.0397, shift=60, 95% CI=0-150). However, after the inclusion of total household assets (the value of household-owned livestock, equipment, household possessions, any land as well as individual income from employment), no significant difference could be detected between the distributions of pay (Wilcoxon rank sum, W=35612, p-value = 0.5263).

	Northern	Brong Ahafo	(all)	Northern	Brong Ahafo	(all)
Economic Sector	**Median pay**			**# respondents**		
Agriculutre	400	145	250	202	199	401
Fishing	1400	3,000	1,500	30	1	31
Mining and quarrying	NA	660	660	0	6	6
Manufacturing	550	750	635	8	8	16
Construction	NA	3,900	3,900	0	2	2
Transportation/storage/comm.	NA	3,672	3,672	0	4	4
Finance/insurance/services	3000	315	660	4	4	8
Utilities	120	NA	120	1	0	1
Wholesale/retail trade	600	490	600	31	20	51
Community/social services	365	2,250	1,115	2	2	4
Other (specify)	2040	2,940	2,880	3	2	5
(all)	**500**	**200**	**400**	**281**	**248**	**529**

Table 5-5: Table displaying median pay and the number of people employed by economic sector in the study population.

There was a wide variety of problems experienced by those in employment. The most common problem experienced by persons in relation to their work was financing with nearly half of all problems reported (455/1110) being about finance. The next most reported problem was erratic weather conditions, which fits with most people working in agriculture. Differences between the Brong-Ahafo and Northern regions included more reporting of inadequate fish catch in the Northern region which fits with the greater prevalence of fishermen in that region. A table of all reported problems is shown in **Table 5-6**. Other problems aggregated into the 'Other' column of this table include crop destruction by pests and cattle, bush-fires, poor roads/bad infrastructure, flooding and infertile land. The total number of reported problems by village is shown below in **Table 5-7**. People in Senyase reported more problems in relation to work than anywhere else.

Locality	Finance	Difficult access to land	Erratic weather conditions	Uncertain demand for output	Poor health	Inadequate fish catch	High cost of inputs e.g. fuel	Difficulty in marketing products	Unavailability of inputs	Low price of products	Other (specify)	Total problems
WIAE	42	1	6	5	6	1	21	2	5	3	0	92
ASUBENDE	27	1	11	2	9	0	0	1	2	8	7	68
NTERESO	33	0	16	7	3	1	3	9	4	3	12	91
TAKUMDO	37	7	11	11	3	0	14	5	0	7	0	95
SENYASE	44	3	21	6	7	0	4	0	9	10	10	114
JAGBINGBINDO	30	6	22	7	2	0	7	10	6	7	2	99
BAAYA	48	0	9	0	7	0	1	1	0	12	4	82
CHACHE	31	0	12	7	1	2	10	4	2	5	1	75
OHIAMPE	33	1	10	2	8	0	0	2	0	6	0	62
AGBLEKEME	30	2	9	1	5	9	0	7	9	6	5	83
NSUNUA	33	5	12	7	2	12	14	8	1	7	0	101
NEW LONGORO	42	0	11	7	4	1	1	8	1	5	9	89
KYINGAKROM	25	2	3	6	4	0	2	4	6	2	5	59
Northern	236	21	88	45	22	25	69	45	27	38	20	636
Brong Ahafo	219	7	65	23	39	1	8	16	18	43	35	474
(all)	**455**	**28**	**153**	**68**	**61**	**26**	**77**	**61**	**45**	**81**	**55**	**1110**

Table 5-6: Total number of problems related to work experienced by people across different localities.
The most common problem was finance, followed by erratic weather.

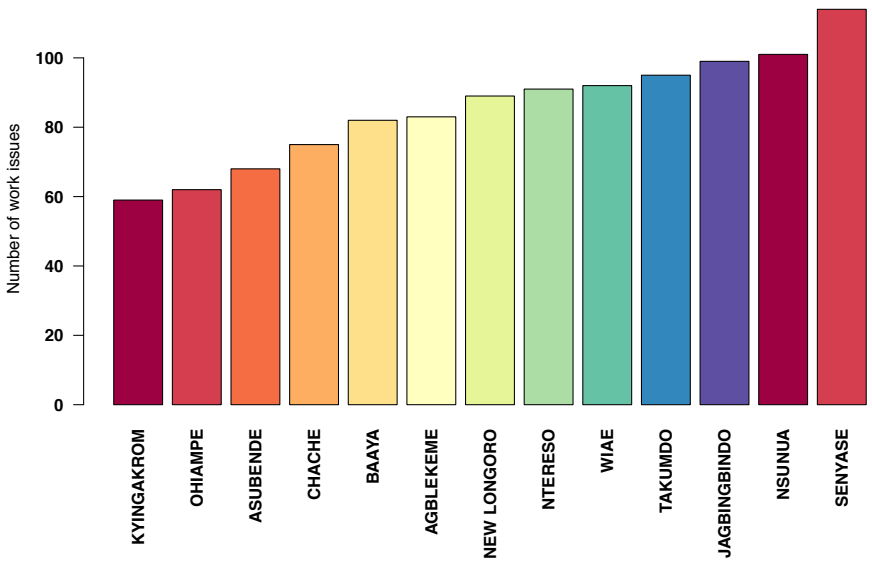


Figure 5-7: Total number of problems listed in relation to work, sorted by village.
People from Senyase reported the most difficulties in relation to their work.

Most farmers appeared to work as subsistence farmers, earning little or no reported income from their crops. Households were asked extensive questions about both the equipment/tools and the household possessions they owned. They were also asked to report the current value of each of these things. The average values for the most common tools/equipment are displayed in **Table 5-7**. Across the 280 households in the study, the total value of tools and equipment reported was GH₵ 36,132. Many households often had access to only rudimentary farming tools such as axes, cutlasses and hoes, and the average value of these common tools was relatively constant across regions. A smaller, but significant number of farmers also used pesticide/insecticide spraying pumps, but this appeared to have no effect on farmers' income (Wilcoxon rank sum, W = 3698, p-value = 0.4186).

Persons involved in the fishing industry also owned canoes and nets, which had relatively high intrinsic value compared to rudimentary farm tools. Persons reporting fishing as their main job, were more common in the Northern Region (n=30) than the Brong-Ahafo region (n=1), and this was reflected in the differing values placed on canoes in the study. The average value placed on canoes in the Northern Region was GH₵ 286, whilst in Brong Ahafo it was GH₵ 158. Fishing appeared more profitable than agriculture/farming, but the term for agriculture and fishing is sometimes used interchangeably, as many fishermen also tend to grow their own subsistence crops as reflected by the fact that more canoes were actually owned by households in Brong Ahafo (n=31) than in the Northern region (n=18). The distribution and total value of tools and equipment by village and region are shown in **Table 5-8**. Wiae, Baaya and Ohiampe had the most wealth in terms of tools and equipment compared with other villages, and these were the villages where the greatest number of canoes were recorded.

Locality	Hoe	Axe	Rake	Shovel	Pick-axe	Sickle / reaping hook	Cutlass	Trailer / cart	Spraying machine	Outboard motor	Canoe	Other
WIAE	3.95	5	5	5	9	2	5		28.25		172	57
ASUBENDE	2.47	3				2	4		26.67			10
NTERESO	3.95	6	7	10	10	3	6		41.25	200	140	13
TAKUMDO	2.73	4		11	9		6		29.25		120	28
SENYASE	4.59	4			10	2	6	13.5	46.66		250	4
JAGBINGBINDO	3.11	3		8	8		5	6	20		100	1,500
BAAYA	5.41	5				9	6		30		276	22
CHACHE	4.31	4			7	4	6		49.17			
OHIAMPE	5.8	9		12		8	8		40	1,500	325	99
AGBLEKEME	3.49	7	4	8	9		5		29.62			
NSUNUA	2.76	4		0	4	1	4	15.5	23.09			
NEW LONGORO	2.41	3			9	1	4		26			5
KYINGAKROM	4.42	4		7	7	7	7		30		200	4
Northern	**3.52**	**4.97**	**5**	**7.41**	**7.28**	**2.29**	**5**	**12.33**	**31.84**	**200**	**157.69**	**184**
Brong Ahafo	**3.89**	**4.55**		**9.5**	**8.2**	**6.18**	**5.43**	**13.5**	**32.7**	**1500**	**285.75**	**37**
(all)	**3.80**	**4.71**	**5.33**	**7.59**	**8.06**	**3.87**	**5.32**	**11.67**	**32.30**	**850.00**	**197.93**	**174.19**

Table 5-7: Table of average value assigned to tools and equipment, shown by village and region.
The final column 'Other', includes items such as guns, fishing nets and plastic chairs.

Locality	Hoe	Axe	Rake	Shovel	Pick-axe	Sickle / reaping hook	Cutlass	Trailer / cart	Spraying machine	Canoe	Other	TOTAL GHC
WIAE	100	4	1	6	4	1	72		4	14	83	9,408
OHIAMPE	129	16		2		35	84		13	13	18	8,519
BAAYA	70	8				8	47		4	15	39	6,408
JAGBINGBINDO	46	5		9	2		31	2	1	1	1	2,017
SENYASE	110	16			1	2	67	2	4	2	3	1,752
NTERESO	73	12	1	1	1	1	44		5	2	13	1,444
CHACHE	117	16			2	4	69		8			1,397
TAKUMDO	33	6		4	1		24		4	1	41	1,266
NSUNUA	124	22		1	5	4	71	6	13			1,067
NEW LONGORO	185	17			3	1	71		6		4	961
KYINGAKROM	64	7		1	3	4	40		3	1	2	907
AGBLEKEME	41	11	3	4	5		37		5			616
ASUBENDE	66	5				6	32		3		1	370
Brong Ahafo	**624**	**69**		**3**	**7**	**56**	**341**	**2**	**33**	**31**	**67**	**18,917**
Northern	**534**	**76**	**5**	**25**	**20**	**10**	**348**	**8**	**40**	**18**	**138**	**17,215**
(all)	**1158**	**145**	**5**	**28**	**27**	**66**	**689**	**10**	**73**	**49**	**205**	**36,132**

Table 5-8: Table of distribution of farming tools across villages and regions.
The last column gives the total value of equipment within each locality. The rows of the table are ordered by total value reported in each locality.

Detailed information on household possesions were also collected, including any buildings or plots of land that were owned. The total value of reported possesions in this study was GH₵ 244,329, with a median household value for possessions of GH₵ of 355. Most households had access to basic household equipment, such as cooking utensils, furniture, torches and buckets (used for washing and fetching drinking water). The majority of households also owned at least one bicycle (69.3%), a radio (65.7%), and a large proportion also owned at least one mobile phone (40%). No households had access to a landline. The average value of common categories of possessions is shown in **Table 5-9**, stratified by village and by region. Less common items that are not included in **Table 5-9** included refridgerators (New Longoro and Kyingakrom), a laptop computer (Chache), washing machines (New Longoro and Agblekeme) and cooking stoves. Some households owned holdings of non-farm land (20.0%).

Many households included at least one child at school (67.1%). And ownership of textbooks was high. Ntereso was the only community in which no ownership of textbooks was reported. Most electrical items were found in the communities with electrical infrastrucutre (New Longoro and Agblekeme). However, a low number of households in communities without regular access to electricity powered electrical items with petrol generators.

Locality	Room Furniture	Sewing Machine	Radio	Radio casette player	Bicycle	Motorcycle	Book (textbooks)	Cellphone	TV	Lanterns / gas lights	Buckets	Cutlery / Utensils	Cups and plates	Torches	Land (non-farm)	Other (specify)
WIAE	95.87	67	14	77	76	738	22	40.56	140	5	5	98.06	2.81	3	1,130	5
ASUBENDE	17.39	83	10	39	53		4	23.75		7	4	39.68	5.25	2	203	6
NTERESO	112.07	73	18	51	87	1,300		77.86	150	7	6	90.86	9.32	13	840	
TAKUMDO	56.2	60	15	42	79	900	2	45		4	6	58.97	1.98	3	250	3
SENYASE	17.37		12	15	67	500	4	37	100	6	4	19.38	2.52	2	3,500	
JAGBINGBINDO	59.49	102	13	48	71	650	1	62		5	5	48.59	2.18	3	2,563	7
BAAYA	51.28	150	12	38	60	600	4	32.83		8	218	41.31	5.97	3		15
CHACHE	39.68	53	10	35	66	800	2	48.17	225	3	3	64.11	3.03	3	400	303
OHIAMPE	36	94	9	27	58	471	4	37.29		4	6	35.36	2.75	3	33	56
AGBLEKEME	111.44	35	18	78	69	100	4	41.19	75	10	84	22.94	4.77	26	588	20
NSUNUA	64.08	96	16	35	70	825	8	55.56		7	7	35.54	2.38	4	150	
NEW LONGORO	56.69	73	15	68	51	2,000	5	43.85	88.85	6	6	36.17	4.76	10	818	102
KYINGAKROM	79.89	55	9		53		3	46.88	73.75	6	7	23.28	3.51	3	175	18
Northern	**77.42**	**72.5**	**15.65**	**51.23**	**73.96**	**880.26**	**8.79**	**51.61**	**147.5**	**5.56**	**16.69**	**57.4**	**3.9**	**8.22**	**850.15**	**57**
Brong Ahafo	**46.58**	**90.52**	**11.82**	**41.36**	**56.38**	**636.36**	**4.15**	**40.62**	**86.11**	**6.03**	**36.89**	**32.95**	**4.18**	**4.65**	**739.78**	**48**
(all)	**62.21**	**82.36**	**13.79**	**48.04**	**65.81**	**825.51**	**6.17**	**45.72**	**97.27**	**5.78**	**26.94**	**43.71**	**4.05**	**6.46**	**804.82**	**51.44**

Table 5-9: Table of average value assigned to household possessions and other assets, shown by village and region.
Only the most popular categories are shown.

Locality	Room Furniture	Sewing Machine	Radio	Radio casette player	Bicycle	Motorcycle	Book (textbooks)	Cellphone	TV	Lanterns / gas lights	Buckets	Cutlery / Utensils	Cups and plates	Torches	Land (non-farm)	Total Value
NEW LONGORO	105	9	33	6	41	1	94	44	13	15	117	427	344	62	10	36,997
NTERESO	46	4	29	10	22	13		9	1	12	53	84	132	49	15	36,360
AGBLEKEME	80	2	32	12	26	3	143	28	1	15	49	264	255	85	13	24,673
JAGBINGBINDO	56	4	7	8	26	9	23	6		49	35	97	72	44	4	24,032
NSUNUA	25	4	25	5	22	7	35	10		7	42	77	228	47	1	23,774
WIAE	33	7	9	3	30	4	144	13	1	16	29	86	60	32	5	19,921
BAAYA	51	7	17	10	35	2	18	6		14	50	163	108	67		18,642
SENYASE	36		12	2	19	1	17	5	1	7	31	182	117	46	2	12,312
CHACHE	22	4	12	11	21	3	30	9	2	9	26	76	72	32	5	11,954
TAKUMDO	30	1	16	3	25	4	81	7		17	41	53	80	38	8	11,326
OHIAMPE	36	8	11	3	24	9	26	11		6	45	128	100	41	4	9,798
KYINGAKROM	19	6	12		15		60	16	4	8	43	150	158	27	2	9,300
ASUBENDE	14	4	11	3	14		23	4		9	27	123	84	26	9	5,240
NORTHERN	**292**	**26**	**130**	**52**	**172**	**43**	**456**	**82**	**5**	**125**	**275**	**737**	**899**	**327**	**51**	**152,040**
BRONG-AHAFO	**261**	**34**	**96**	**24**	**148**	**13**	**238**	**86**	**18**	**59**	**313**	**1173**	**911**	**269**	**27**	**92,289**
(all)	**553**	**60**	**226**	**76**	**320**	**56**	**694**	**168**	**23**	**184**	**588**	**1910**	**1810**	**596**	**78**	**244,329**

Table 5-10: Number of possessions owned by village and region, for the most popular categories.
Categories not shown include small numbers of household electrical items such as refrigerators (n=5) and electronic devices such as cameras (n=2).

Of the 706 people not of working age (defined as under 15 and over 64), 80 (11.33%) reported as being in work, with 41 of these being under the age of 15. The mean age of those working, who were under 15 was 10 years and 3 months, with an average annual income of 39 GH¢ . Additionally, there was one person whose age was not recorded and they also identified themselves as employed or in work.

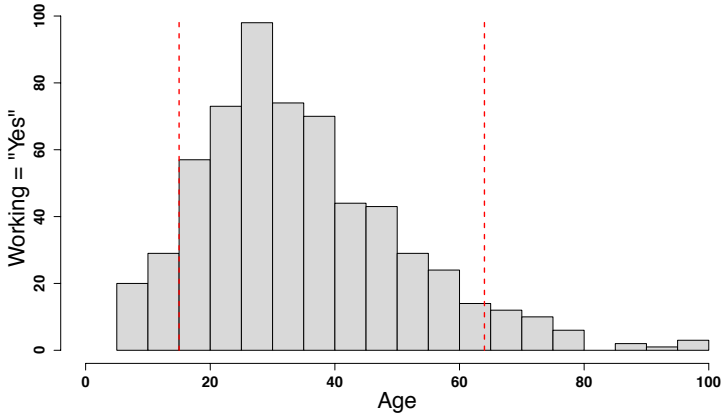


Figure 5-8: Age distribution employed persons.
Histogram displays those responses for those identifying themselves as currently being in employment or working as an unpaid family member (e.g. on a family farm). The dashed red lines indicate the years 15 and 64 which are defined as the lower and upper limits.

5.4.5 Animal Husbandry

Animal husbandry was common across the study population. There were 210 households which kept livestock and these were split mostly evenly between Brong-Ahafo (n=102) and the Northern region (n=108). Many houses kept low numbers of chicken and goats. In Ntereso in the Northern Region there was a professional farmer who had 1,000 chickens, 35 sheep and 50 goats, for a combined livestock value of GH¢ 24,500 which was nearly 70 times the median value of livestock kept by each household (GH¢ 363).

Livestock distribution is show in **Table 5-11**. All respondents were asked to place a total value on each of the categories of livestock they had. **Table 5-9** shows, for each type of animal, the distribution of average values per animal. The average value per animal placed on livestock between the Northern

and Brong-Ahafo regions was broadly similar, indicating that study participants had a good awareness and knowledge about the market value of their animals.

Locailty	Cattle	Chickens	Goats	Pigs	Sheep	Others	(all)
AGBLEKEME	3	173	158	70	112	50	566
ASUBENDE	0	153	28	0	21	27	229
BAAYA	125	215	115	39	14	38	546
CHACHE	59	95	27	0	19	95	295
JAGBINGBINDO	3	205	56	7	41	5	317
KYINGAKROM	3	157	28	0	13	19	220
NEW LONGORO	4	235	102	0	41	37	425
NSUNUA	70	106	93	15	36	10	330
NTERESO	28	1,767	182	0	133	293	2,404
OHIAMPE	21	173	68	0	23	28	313
SENYASE	0	134	70	18	3	10	235
TAKUMDO	1	159	71	17	12	13	273
WIAE	0	97	54	7	40	6	204
Northern	**164**	**2,602**	**641**	**116**	**393**	**472**	**4,389**
Brong Ahafo	**153**	**1,067**	**411**	**57**	**115**	**159**	**1,968**
(all)	**317**	**3669**	**1052**	**173**	**508**	**631**	**6,357**

Table 5-11: Number and type of animals kept by household across the study region.
The livestock listed in the 'Others' column mainly consists of Guinea Fowl and Ducks.

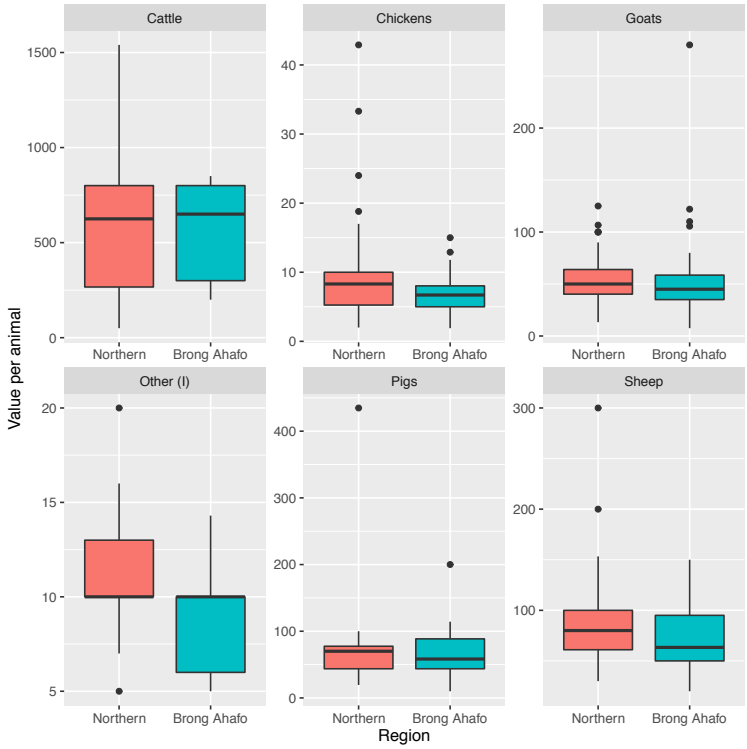


Figure 5-9: Box-plots showing the distribution of values of different types of livestock (per animal).
The plots are grouped by region. The value of each type of livestock was broadly similar between regions. The plot labelled 'Other (I)' refers entirely to fowl and ducks.

Most households with livestock allowed their livestock to roam free-range during the daytime, but there were notable differences in the way animals were housed at night: 68 households stabled animals inside the house; 83 households kept animals outside the house; 45 households kept animals in purpose-built animal sheds at night (also outside the house). Living in a household that kept animals inside the house at night was associated with a lower risk of reporting symptoms of onchocerciasis (χ^2 5.87, OR 0.69, 95% CI 0.50-0.93, p-value 0.015), but an increased risk of having suffered illness or injury in the last 4 weeks (χ^2 6.05, OR 1.42, 95% CI 1.07-1.87, p-value 0.014).

5.4.6 General health and education

Respondents were asked about their health within the last 4 weeks, as well as their health within the last 12 months. Within the last 4 weeks, 285 people (20.6%) had suffered illness, 21 had suffered injury, and one person had suffered both. The most common cause of recent ill-health was non-specific fever (10.4%), followed by cough/cold (1.74%) and non-specific body pains (1.81%). Those aged 35 years and older were nearly twice as likely to suffer ill-health than those under 35 (OR 1.91, 95% CI 1.56-2.33, p-value <0.001). During the last 12 months, 61 people were hospitalized (4.4%) resulting in 393 days of lost productivity. More person-days were lost to hospitalization in the Northern Region than in Brong-Ahafo (242 versus 151).

Of the study population, 1,238 were at least of school age (3 years and above). Of these, nearly 60% reported as ever having attended school (n=740) with 499 currently in school (67.1% of households had at least one child at school). There were 239 persons who were no longer attending school. The exit grade achieved by those persons is shown by a bar plot in **Figure 5-10**. Most students successfully completed their primary school basic education (ages 6 - 12, grades P1-P6) with 61.2% indicating successful completion of P6. Some of the school grades on this bar plot are shown together due to a reform of the education system which was undertaken in Ghana in 1987 (*268*).

The previous middle school system of 4 years (M1-4), followed by a senior school of 5 years (S1-5) was replaced by the junior school system (ages 12-15; JSS1-3) and the senior school system (ages 15-18; SSS1-3) respectively, each of which last 3 years. The programme of Basic Education for 6-15 year olds is therefore completed at the end of JSS3, with entry to SSS1-3 being via successful completion of the Basic Education Certificate exams. 43.5% of school leavers completed JSS3 (or the old M3), but the proportion of persons continuing to senior secondary school drops to just 18.1%.

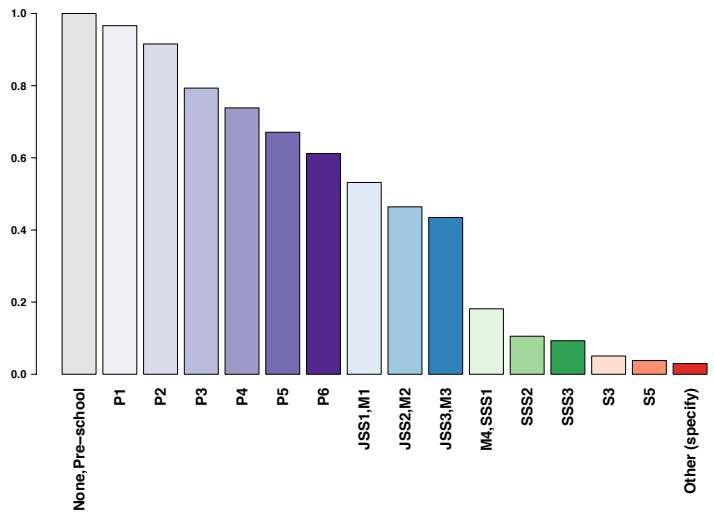


Figure 5-10: Count of persons successfully completing each school year.
JSS3 is the exit of the current Basic Education programme in Ghana and is a natural end-point for many students who do not go on to enter the senior school system.

Of the 239 school leavers 80 (33.5%) obtained either their Middle School Leavers Certificate under the old system (n=19) or the newer Basic Education Certificate Examination (n=61). A further 12 (5.0%) gained Senior School Certificate Examinations and 7 (2.9%) took either professional, technical or teacher training exams. More than half of school leavers (58.6%, n=140), left with no qualifications at all. The median length of time taken to get to school in either region was 10 minutes, with most (91.6%) walking to school.

Qualification	Northern	Brong Ahafo	(all)
None	64	76	140
MSLC	11	8	19
BECE	24	37	61
Teacher Training	2	2	4
SSCE	5	7	12
Tech/Prof. Cert.	1	0	1
Tech/Prof. Dip.	0	2	2
(all)	**107**	**132**	**239**

Table 5-12: Number and type of qualifications achieved by school leavers, stratified by study region.

5.4.7 Household characteristics

Information was collected on a range of household characteristics, including the households' community engagement, the main respondents' perception of other social groups, the type of materials used in household construction, the number of rooms, ownership status, the cooking facilities, access to services and utilities, the upkeep and state of the household and immediate surroundings, access to water supplies for drinking and general use and the type of toilet facilities used by the household.

There was a marked difference in how interview respondents perceived different social groups. Most respondents had a positive reaction ('I trust them a lot') to their relatives (86.1% of responses). Community distributors were also a highly-trusted group with 72.9% of responses positive. The least trusted groups (response of 'Not at all') were people from different tribes, different regions, and people from other African countries, with just 16.1% of households stating they trust this group a lot. **Figure 5-11** shows the responses given for each socioeconomic or demographic group.

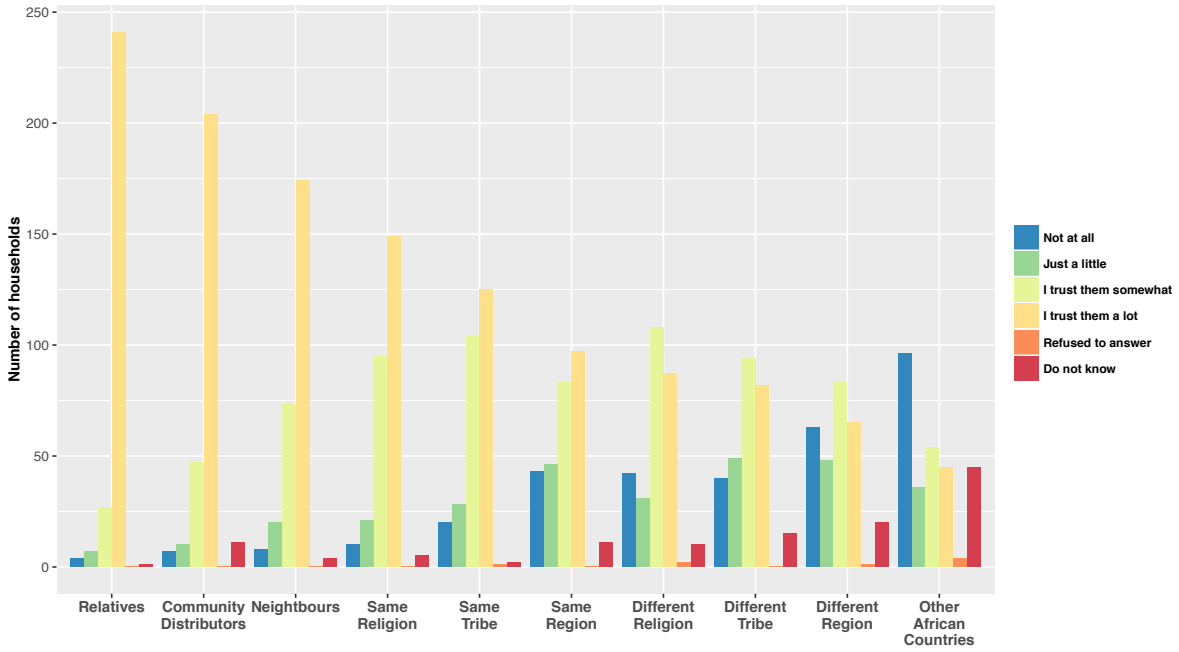


Figure 5-11: Distribution of responses about how the main survey respondent felt about different social groups.

The most marked difference in household characteristics across communities was in the source of water for drinking and general use (e.g. cooking and washing). Most communities relied solely on either river water, or on using a borehole. In Agblekeme, Chache and Takumdo there was mixed access to one of these two water sources, but in each case one water source was favoured over the other in each community. Unexpectedly using the river as the drinking water source was significantly associated with a slightly reduced odds of reporting suffering onchocerciasis (χ^2 7.591, OR 0.70, 95% CI 0.54-0.90, p-value 0.006).

Village	Are you happy with your life	Electricity in household	Household keeps animals	Household has no toilet	Drinking water from river	Drinking water from borehole	Household treats drinking water	General-use water from river	Access to mobile phone	Access to internet	Access to computer	Access to paid TV services
AGBLEKEME	17	0	19	19	4	15	0	4	16	1	0	0
ASUBENDE	14	0	13	17	18	0	4	18	11	0	0	0
BAAYA	15	0	18	19	20	0	5	20	16	0	0	0
CHACHE	16	1	11	20	5	14	2	9	9	0	0	0
JAGBINGBINDO	16	0	19	20	20	0	4	20	10	0	0	0
KYINGAKROM	17	7	12	10	0	20	0	0	18	2	1	0
NEW LONGORO	36	23	23	11	0	37	0	1	36	0	0	1
NSUNUA	17	0	13	19	20	0	1	20	14	0	0	0
NTERESO	18	0	19	16	20	0	2	20	12	0	0	0
OHIAMPE	18	0	18	11	0	20	0	0	16	0	0	0
SENYASE	21	0	18	22	22	0	8	22	13	0	0	0
TAKUMDO	17	0	13	20	16	4	4	17	10	0	0	0
WIAE	15	0	14	10	0	20	1	3	11	0	0	0
Northern	**116**	**1**	**108**	**124**	**85**	**53**	**14**	**93**	**82**	**1**	**0**	**0**
Brong Ahafo	**121**	**30**	**102**	**90**	**60**	**77**	**17**	**61**	**110**	**2**	**1**	**1**
(all)	**237**	**31**	**210**	**214**	**145**	**130**	**31**	**154**	**192**	**3**	**1**	**1**

Table 5-13: Distribution of various household characteristics related to socioeconomic setting and access to services.
Numbers show the total number of households in each locality with e.g. access to electricity, or who use river water as their main drinking source.

5.4.8 Health with respect to onchocerciasis

There was wide variation in prevalence of symptoms associated with onchocerciasis between villages and between regions. There were 91 households (32.5%) with at least one person suffering symptoms of onchocerciasis (self-reported). The average number of persons suffering onchocerciasis, per household with at least one infected individual was 3.36, the median was 2 persons, and the range was from 1 to 12 infected persons within a single household. The distribution of self-reported cases per household with at least one case, is shown in **Figure 5-12** below.

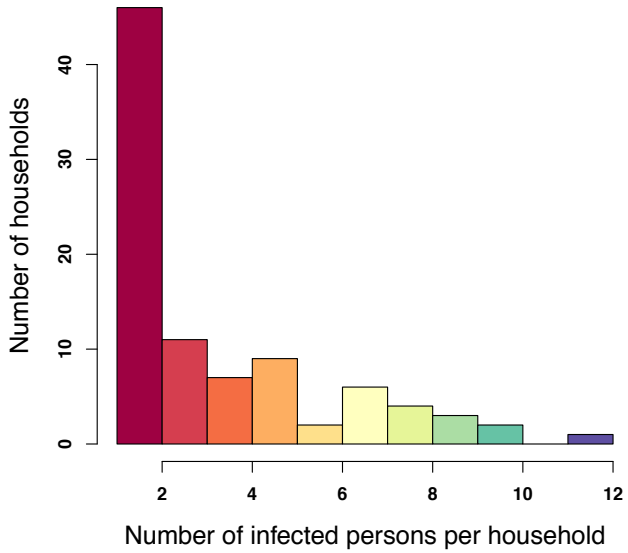


Figure 5-12: Distribution of number of persons per household who report suffering onchocerciasis, where at least one person reports they suffer from onchocerciasis.

The prevalence of self-reported cases of onchocerciasis ('Did you suffer onchocerciasis in the last 12 months?') was 22.1% (n=306). People who reported that they suffered onchocerciasis in the last 12 months were asked about their symptoms. The overall prevalence of blindness in the study region was 1.8% (n=25), troublesome itch was 17.9% (n=248) and skin disease was 15.6% (n=216).

Table 5-14 shows the prevalence of symptoms across the two regions and across villages. Within the study population there were four times as many blind persons (n=20, 3.0%) living in the Brong-Ahafo region than the Northern region (n=5, 0.7%), and skin disease (25.2%) and troublesome itch (27.5%)

were also more prevalent in Brong-Ahafo than the Northern region (6.6% and 9.0% respectively). Asubende had the highest prevalence of blindness at 11.0%, whilst nearby Ohiampe had the highest prevalence (self-reported) skin disease (34.0%). Residents of Agblekeme reported suffering troublesome itch (self-reported) more than any other village (44.0%).

Locality	Did you suffer onchocerciasis?	Blindness	Other Symptoms	Skin Disease	Troublesome Itch
AGBLEKEME	0.5	0.000	0.010	0.290	0.440
ASUBENDE	0.333	0.110	0.000	0.230	0.250
BAAYA	0.284	0.000	0.000	0.280	0.280
CHACHE	0.042	0.010	0.010	0.010	0.010
JAGBINGBINDO	0.038	0.000	0.030	0.020	0.000
KYINGAKROM	0.25	0.030	0.010	0.210	0.210
NEW LONGORO	0.288	0.030	0.010	0.210	0.240
NSUNUA	0.076	0.000	0.000	0.030	0.050
NTERESO	0.041	0.030	0.010	0.000	0.000
OHIAMPE	0.365	0.030	0.000	0.340	0.330
SENYASE	0.327	0.000	0.010	0.250	0.310
TAKUMDO	0.134	0.000	0.030	0.070	0.080
WIAE	0.162	0.020	0.030	0.060	0.080
Northern	**0.139**	**0.007**	**0.020**	**0.066**	**0.090**
Brong Ahafo	**0.309**	**0.030**	**0.006**	**0.252**	**0.275**
(all)	**0.22**	**0.018**	**0.013**	**0.156**	**0.179**

Table 5-14: Distribution of self-reported prevalence of sysmtpoms related to infection with onchocerciasis. Prevalences are crude prevalence across the total study population interviewed within each locality.

5.4.9 Defining compliance

Individual and household level variables were investigated for a statistical association with onchocerciasis infection and with compliance to community-directed treatment with ivermectin. Compliance was measured in 3 different ways. Initially I split the population into compliers versus non-compliers based on their response to the question 'Did you take ivermectin in the last 12 months?'. I was further able to assign persons into categories of compliance based on their response to the question 'Why did you not take ivermectin in the last 12 months?'. Persons were either classified as: a) 'Good'; persons who were compliant (those who did take ivermectin in the last 12

months) b) 'Random'; persons who did not take ivermectin in the last 12 months, due to external factors, rather than internal value judgements (e.g. pregnancy, under height/weight, absent during distribution) and c) 'Systematic'; persons who did not take ivermectin, and their reasons for not taking it indicated they had chosen to not comply (e.g. 'Drug not useful', 'Previous side-effects', 'Refused' and 'Not suffering from oncho'). The distribution of types of compliers based on this classification system are shown in **Table 5-15**.

Did you suffer from onchocerciasis in the last 12 months	Compliance Type			
	Good (#)	Random (#)	Systematic (#)	(all)
Yes	0.863 (264)	0.124 (38)	0.013 (4)	306
No	0.725 (774)	0.229 (244)	0.046 (49)	1067
(all)	**1038**	**282**	**53**	**1373**

Table 5-15: Types of complier stratified by self-reported onchocerciasis infection status.
The proportions in the table are marginalised by infection status (row proportions sum to 1).

A binary logistic regression revealed that short-term compliance status was significantly associated with onchocerciasis status, with the odds of reporting suffering onchocerciasis lowest in the systematic non-compliers group, as seen in **Table 5-16**.

	Odds Ratio	95% CI	Z value	p-value
Random:Good	0.457	0.312 - 0.653	-4.161	< 0.0001
Systematic:Good	0.239	0.072 – 0.594	-2.724	0.0064

Table 5-16: Results of a binary logistic regression of compliance status against self-reported infection status.

The distribution and proportion of compliance types is shown in Table 1-17. Chache, Kyingakrom, New Longoro and Nsunua were all identified as containing individuals who were classed as systematic non-compliers. The overall level of compliance in Chache was particularly poor with just 32.3% of people complying with CDTI last year.

	Compliance Type		
Village	Good	Random	Systematic
AGBLEKEME	0.791	0.209	0.000
ASUBENDE	0.828	0.172	0.000
BAAYA	0.866	0.134	0.000
CHACHE	0.323	0.427	**0.250**
JAGBINGBINDO	0.769	0.231	0.000
KYINGAKROM	0.838	0.103	**0.059**
NEW LONGORO	0.718	0.218	**0.064**
NSUNUA	0.638	0.219	**0.143**
NTERESO	0.671	0.329	0.000
OHIAMPE	0.843	0.157	0.000
SENYASE	0.836	0.164	0.000
TAKUMDO	0.798	0.202	0.000
WIAE	0.862	0.138	0.000

Table 5-17: Proportion of persons stratified by compliance type within each village.

A continuous variable for longer-term compliance recall was constructed. This was defined as the difference between the most recent time someone took ivermectin and the first time they took it, minus the number of treatments they said they had received, with an adjustment for multiple treatment rounds in 2010 in all villages except Nsunua and Chache. Values under 0 represent an individual either overestimating the number of treatments they took or multiple years of treatment rounds. Values over 0 represent individuals either underestimating the number of treatments they took, or they are potential missed treatment rounds. This distribution had a median and mode of 0 and a range of -10 to 14 (**Figure 5-13**). This distribution was divided into good and poor compliers, with the cut-off for good compliers chosen such that the ratio of good to poor compliers was as close as possible to the ratio of those who took ivermectin in the last 12 months to those who did not. Persons with a missed treatments score of less than or equal to 2 were defined as good compliers, with a score above two being classified as a poor responder.

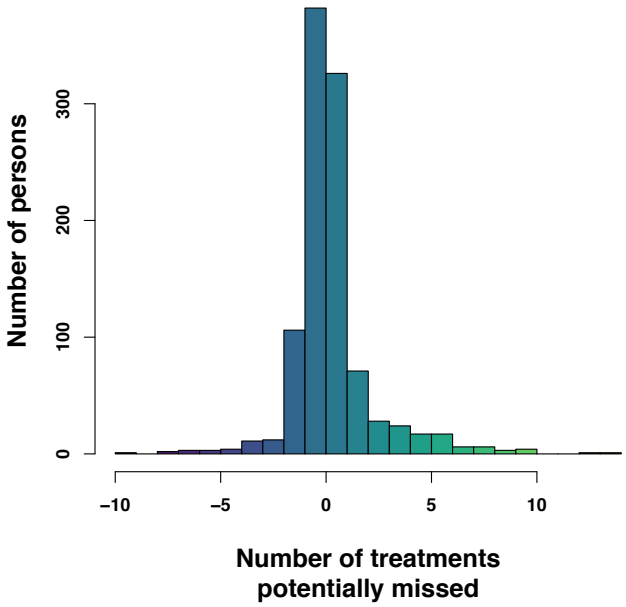


Figure 5-13: The number of treatments missed by an individual according to their recall of when they last took ivermectin, the first time they took ivermectin and the number of treatments in total that they received.

It was not possible to confirm treatment compliance with community treatment registers. These were not generally available in the communities at the time of the surveys, and in some cases, when available, the data were missing or incomplete (see **Figure 5-14**). I used data from the Ghana Health Service to ascertain coverage in the study communities where possible, however comparing these data to some of the few available treatment summaries in community treatment registers, there were discrepancies between coverage in the treatment registers and that obtained from the Ghana Health Service.

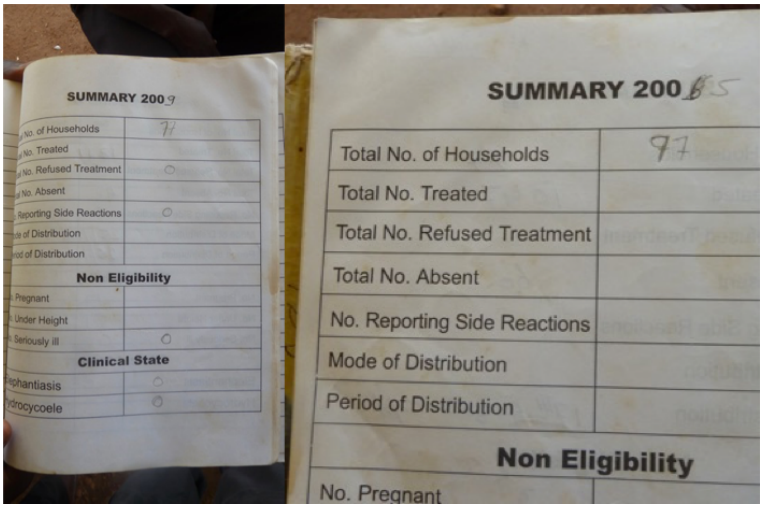


Figure 5-14: An example of some of the treatment records that were recovered whilst on site.
These records for a community in the Northern region were incomplete, and records were often not available at the time of the survey.

Ascertaining an accurate estimate of the actual number of treatments taken by an individual was extremely difficult. For all persons aged 29 and over (those who would have been eligible for treatment at the time the first distributions took place in Ghana), who were born in their village (remove effect of migration) I compared the level of compliance from **Figure 5-13**, against a level of compliance derived from subtracting the recalled number of rounds of treatment taken, with the number of rounds of treatment provided in the village (estimated from treatment data from Ghana Health Service, the Onchocerciasis Control Programme Epidemiological Database and (*113*), see **Table 5-18**).

The distributions for these two groups were similar, suggesting that respondent recall of ivermectin treatment history could be useful, however it was thought that accurate recall of up to 25 rounds of treatment with ivermectin (over 24 years in some cases) would be too prone to recall bias.

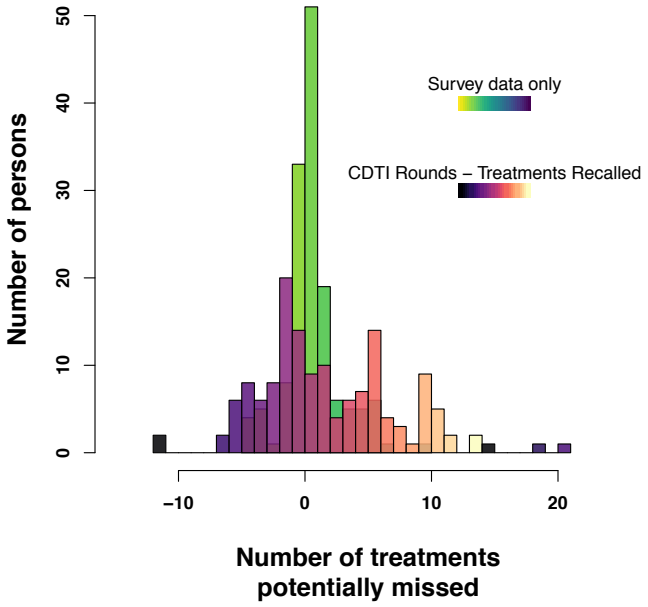


Figure 5-15: Comparison of pure respondent recall of number of treatments they received, vs, the number of treatments actually administered to each village.

5.4.10 Variables associated with onchocerciasis status

Bivariate analyses were carried out, using binary logistic regression to test the association of sociodemographic characteristics with onchocerciasis status. Odds ratios and 95% confidence intervals (CI) were calculated to measure the strength of association with each variable. Both individual level and household level characteristics were assessed. The outcome variable was 'Did you suffer from onchocerciasis in the last 12 months?' with a 'Yes/No' response. **Table 5-19** displays the individual level factors that were tested for an association with onchocerciasis.

Village	IVM START	Estimated number of rounds	Missed years	Semi-annual treatment?	Distributor payment GH¢	Paid by	Distributor religion	Number of distributors	Political views	Community choose distributor
AGBLEKEME	1987	23	2005	2,009	5	GHS	Christian	3	Yes	Yes
ASUBENDE	1987	25	2005	2,009	10	GHS	Christian	1	No	Yes
BAAYA	1987	24	2005	2,009	10	GHS	Christian	2	Yes	Yes
CHACHE	1997	11	2007-2009	No	DK	NA	Muslim	2	No	No
JAGBINGBINDO	1988	16	2007-2008	2,010	10	GHS	Traditional	1	Yes	Yes
KYINGAKROM	1989	21	2005-2008	2,009	NA	NA	Christian	1	Yes	Yes
NEW LONGORO	1989	22	2005-2007	2,009	7	GHS	Muslim	2	Yes	Yes
NSUNUA	1991	16	DK	No	5	GHS	Christian	3	No	No
NTERESO	1997	11	DK	2,010	0	NA	Christian	2	Yes	Yes
OHIAMPE	1988	25	2005	2,009	12	GHS	Christian	2	Yes	Yes
SENYASE	1988	25	2005	2,009	DK	NA	Christian	1	Yes	No
TAKUMDO	1991	19	2007-2009	2,010	15	GHS	Traditional	4	No	Yes
WIAE	1991	14	2008	2,010	15	GHS	Traditional	2	No	Yes

Table 5-18: Village knowledge about CDTI interventions.
The start date of ivermectin treatment is estimated from several data sources (see text).

Age was positively associated with onchocerciasis infection as was expected, with a prevalence of 53.4% in the over 35s. Other factors that had significant positive associations (at $p < 0.05$) with self-reported infection with onchocerciasis were being able to do written calculations, suffering illness or injury in the last 4 weeks. Factors which were negatively correlated (i.e. exposure reduces risk of reporting infection with onchocerciasis) were having taken ivermectin in the last 12 months, facing erratic weather as a problem at work (chosen as a proxy for working outdoors), and having attended the last village meeting (restricted to those aged 12 years and above).

Individual: Did you suffer onchocerciasis in the last 12 months						
		Yes	No			
Individual Variables	**OR**	**L95%CI**	**U95%CI**	**χ2**	**p-value**	**signif**
Gender	**1.106**	**0.858**	**1.427**	**0.605**	**0.437**	
Female		160	531			
Male		146	536			
Age	**2.246**	**1.679**	**3.004**	**30.707**	**0**	***
Age > 35		95	178			
Age ≤ 35		211	888			
Current marital status	**1.288**	**0.986**	**1.682**	**3.467**	**0.063**	*
Married		111	327			
Currently not married		195	740			
Education level (age > 12)						
None:Primary	**1.339**	**0.759**	**2.360**	**1.017**	**0.313**	
None:Secondary	**0.810**	**0.527**	**1.245**	**0.919**	**0.338**	
None		151	418			
Primary		17	63			
Secondary and senior		37	83			
Employment status (age > 12)	**0.85**	**0.587**	**1.231**	**0.741**	**0.389**	
Employed		151	429			
Not employed		53	128			
Religion	**0.97**	**0.67**	**1.4**	**0.025**	**0.875**	
Muslim		47	136			
Christian		199	559			
Duration of stay in village	**1.276**	**0.822**	**1.979**	**1.187**	**0.276**	
Stay > 10 years		237	743			
Stay ≤ 10 years		28	112			
Monthly average family income	**1.211**	**0.865**	**1.695**	**1.242**	**0.265**	
GH¢ < 600		255	859			
GH¢ ≥ 600		51	208			
Monthly average family income	**0.891**	**0.686**	**1.157**	**0.753**	**0.386**	
GH¢ = 0		187	681			
GH¢ > 0		119	386			
Can read English	**1.241**	**0.913**	**1.687**	**1.909**	**0.167**	
Yes		77	200			
No		201	648			
Can do written calculations	**1.371**	**1.041**	**1.806**	**5.052**	**0.025**	**
Yes		120	303			
No		158	547			
Illness or injury in last 4 weeks?	**1.863**	**1.402**	**2.477**	**18.694**	**<0.001**	****
Yes		96	210			
No		210	856			
Take IVM in last 12 months?	**0.42**	**0.295**	**0.598**	**24.319**	**<0.001**	****
Yes		42	293			
No		264	774			
Erratic weather problems at work	**0.619**	**0.392**	**0.976**	**4.331**	**0.037**	**
Yes		24	129			
No		282	938			
Attended village meeting (age > 12)	**0.4**	**0.275**	**0.581**	**23.906**	**<0.001**	****
Yes		47	204			
No		155	269			

Table 5-19: Individual level demographic factors and onchocerciasis status.
OR=Odds Ratio, signif= significance level; * p < 0.1, ** <0.05, *** <0.005, **** <0.001.

Current marital status had a slight association, however this was tested across all ages. If the association with being married was restricted to all persons aged 12 and over, there was no statistical association between being currently married and suffering onchocerciasis (OR=0.836, 95% CI=0.608-1.15, χ^2=1.20, p-value=0.272). An individual's educational status, current employment status or religion were not found to be significant. Interestingly, the length of stay of 10 years or more in the village of residence was also not associated with suffering onchocerciasis.

At the household level (**Table 5-20**), there was many factors which were associated with onchocerciasis status. The first question all main respondents were asked was 'Are you happy with your life?'. Individuals who were living in household with a positive response to this question were nearly 3 times as likely to report problems with onchocerciasis (p-value <0.001). Other household-level variables positively correlated with onchocerciasis problems were related to hygiene and health. Factors which were associated with an increased risk of an individual reporting onchocerciasis were living in a household that disposes of rubbish in the surrounding bush; living in a household which has no formal toilet (e.g. pit latrine or other public toilet) and uses the surrounding bush; households where the main drinking water source was a borehole. This was unexpected as the boreholes are located away from rivers where the blackfly breeding grounds are found.

Living in a household which was located close to the main drinking water source (< 200 metres), regardless of type of water source, was strongly positively associated with reporting suffering onchocerciasis. There was a weak association between owning animals in the household and onchocerciasis (p-value < 0.1) but more strikingly, those households who stabled their animals in their home at night had approximately half the odds of reporting onchocerciasis (OR=0.685).

3 variables concerning household construction were assessed. These were all found to be significant, but these could be a proxy for regional or community level effects as the housing in a single community tends to be quite homogenous.

The household view of external parties was investigated for statistical associations. The level of trust for the community distributors had no association with odds of reporting onchocerciasis in this study.

Persons living in households with a positive view of people from different tribes and from other African countries both had a much-reduced odds of reporting symptoms of onchocerciasis.

Figure 5-16: Typical scene of a mudwall home, with thatched roof and external area for cooking. In many cases, livestock are free range and roam freely around villages.

Household: Did you suffer onchocerciasis in the last 12 months						
		Yes	No			
Household Variables	**OR**	**L95%CI**	**U95%CI**	**χ^2**	**p-value**	**signif**
Are you happy with your life?	**2.886**	**1.852**	**4.496**	**23.663**	**<0.001**	********
Yes		282	851			
No		24	209			
Where do you dispose of refuse?	**1.369**	**1.034**	**1.811**	**4.845**	**0.028**	******
In the bush		221	699			
Elsewhere		85	368			
What type of toilet in household	**1.557**	**1.098**	**2.208**	**6.256**	**0.012**	******
None (Bush/beach)		256	833			
Other		45	228			
Main drinking water source	**1.325**	**1.027**	**1.709**	**4.692**	**0.03**	******
Borehole		153	459			
River/Stream		153	608			
Distance to drinking water source	**6.484**	**4.78**	**8.797**	**168.103**	**<0.001**	********
Distance < 200 metres		244	403			
Distance > 200 metres		62	664			
Does your household keep animals?	**1.358**	**0.962**	**1.919**	**3.038**	**0.081**	*****
Yes		259	856			
No		47	211			
Where are animals stabled at night?	**0.685**	**0.504**	**0.931**	**5.867**	**0.015**	*******
Inside the house		64	296			
Outside the house		237	751			
Type of household	**2.732**	**2.1**	**3.555**	**58.318**	**<0.001**	********
Compound house		194	414			
Other		112	653			
Construction material for roof	**1.274**	**0.988**	**1.644**	**3.483**	**0.062**	*****
Corrugated iron sheets		152	464			
Other		154	599			
Construction material of walls	**1.758**	**1.207**	**2.56**	**8.839**	**0.003**	*******
Mud/Mud bricks		269	856			
Other		37	207			
Do you trust the community distributors	**0.878**	**0.663**	**1.163**	**0.819**	**0.365**	
I trust them a lot		216	776			
Other		90	284			
Do you trust people from different tribes	**0.609**	**0.442**	**0.838**	**9.379**	**0.002**	*******
I trust them a lot		56	287			
Other		250	780			
Trust people from other African countries	**0.39**	**0.25**	**0.609**	**18.213**	**<0.001**	********
I trust them a lot		24	191			
Other		282	876			

Table 5-20: Household level demographic factors and onchocerciasis status.
OR=Odds Ratio, signif= significance level; * p < 0.1, ** <0.05, *** <0.005, **** <0.001.

5.4.11 Variables associated with short-term compliance to CDTI

The same bivariate analyses were carried out to test associations between sociodemographic characteristics and short-term compliance to community-directed treatment with ivermectin. Both individual level and household level characteristics were assessed. The outcome variable was 'Did you take ivermectin in the last 12 months?', with a response of 'Yes or No'.

At the individual level, there was several sociodemographic characteristics that were associated with an individuals' recent compliance with CDTI. Being female was negatively associated with compliance in the last year. This was tested across all ages. To test for an effect of pregnancy or breast feeding which would preclude participation in CDTI I tested the association at ages 16 and under, where I found no association (OR= 0.854, 95% CI=0.626-1.165, $\chi 2$=0.993, p-value=0.319). Persons who were currently married were also more likely to be compliant (p-value < 0.001), and this effect was also seen when children under 12 were excluded from the analysis (p-value=0.004). There was a positive association between compliance and individuals with a length of stay in their village of residence of 10 or more years, as well as having ever attended school (in persons' age 12 and above), being able to read English and perform written calculations. Finally, attending the last village assembly meeting was also positively associated with compliance.

Individual factors which were significantly negatively associated with recent compliance included being currently employed, suffering illness or injury in the last 4 weeks, experiencing erratic weather conditions as a problem faced at work, and declaring the river or oceans as one's main place of work. There were fewer household level factors which were found to significantly influence recent adherence to CDTI. The only factor which was found to be significant was living in a household which owned livestock. This was associated with increased odds of being compliant at the last round of treatment.

Individual Compliance: Did you take IVM in the last 12 months?						
		Yes	No			
Individual Variables	**OR**	**L95%CI**	**U95%CI**	**χ2**	**p-value**	**signif**
Gender	**0.718**	**0.561**	**0.919**	**6.927**	**0.008**	**
Female		497	188			
Male		545	148			
Age	**3.709**	**2.425**	**5.673**	**40.854**	**0**	****
Age > 35		247	26			
Age ≤ 35		794	310			
Current marital status	**2.29**	**1.701**	**3.084**	**30.867**	**0**	***
Married		374	66			
Currently not married		668	270			
Education level						
None:Primary	**1.091**	**0.552**	**2.158**	**0.063**	**0.802**	
None:Secondary	**0.868**	**0.463**	**1.626**	**0.196**	**0.658**	
None						
Primary						
Secondary and senior						
Employment status (age > 12)	**0.467**	**0.248**	**0.88**	**5.78**	**0.016**	**
Employed		506	76			
Not employed		171	12			
Religion	**0.84**	**0.58**	**1.23**	**0.822**	**0.365**	
Muslim		138	45			
Christian		599	164			
Duration of stay in village	**2.056**	**1.358**	**3.112**	**11.997**	**0.001**	**
Stay > 10 years		838	145			
Stay ≤ 10 years		104	37			
Monthly average family income	**0.946**	**0.688**	**1.299**	**0.119**	**0.73**	
GH₵ < 600		844	275			
GH₵ ≥ 600		198	61			
Monthly average family income	**1.11**	**0.859**	**1.436**	**0.638**	**0.424**	
GH₵ = 0		388	117			
GH₵ > 0		654	219			
Ever attended school (age >12)	**1.208**	**0.783**	**1.866**	**0.731**	**0.392**	****
Yes		330	41			
No		353	53			
Can read English	**1.422**	**0.899**	**2.251**	**2.285**	**0.131**	
Yes		253	25			
No		747	105			
Can do written calculations	**1.536**	**1.031**	**2.289**	**4.506**	**0.034**	**
Yes		389	38			
No		613	92			
Illness/injury last 4 weeks	**0.687**	**0.518**	**0.912**	**6.799**	**0.009**	**
Yes		214	92			
No		826	244			
Work river/ocean (age > 12)	**0.172**	**0.083**	**0.353**	**28.128**	**<0.001**	****
Yes		20	15			
No		474	61			
Erratic weather at work	**0.556**	**0.34**	**0.911**	**5.547**	**0.019**	**
Yes		120	26			
No		564	68			
Did you attend village meeting?	**1.847**	**1.084**	**3.147**	**5.207**	**0.022**	**
Yes		231	20			
No		369	59			

Table 5-21: Individual level sociodemographic variables and their association with recent adherence to CDTI.

Household Compliance: Did you take IVM in the last 12 months?

Household Variables	OR	Yes L95%CI	No U95%CI	χ2	p-value	signif
Are you happy with your life?	**0.995**	**0.717**	**1.381**	**0.001**	**0.976**	
Yes		859	278			
No		177	57			
Where do you dispose of refuse?	**1.023**	**0.788**	**1.329**	**0.03**	**0.862**	
In the bush		700	224			
Elsewhere		342	112			
What type of toilet in household	**0.954**	**0.7**	**1.299**	**0.091**	**0.764**	
None (Bush/beach)		821	268			
Other		212	66			
Main drinking water source	**0.846**	**0.661**	**1.082**	**1.774**	**0.183**	
Borehole		456	161			
River/Stream		586	175			
Distance to drinking water source	**1.131**	**0.883**	**1.448**	**0.951**	**0.329**	
Distance < 200 metres		497	150			
Distance > 200 metres		545	186			
Does your household keep animals?	**1.546**	**1.15**	**2.079**	**8.389**	**0.004**	***
Yes		862	254			
No		180	82			
Where are animals stabled at night?	**0.831**	**0.63**	**1.097**	**1.709**	**0.191**	
Inside the house		265	95			
Otuside the house		765	228			
Type of household	**1.16**	**0.904**	**1.488**	**1.366**	**0.243**	
Compound house		469	139			
Other		573	197			
Construction material for roof	**1.026**	**0.801**	**1.314**	**0.042**	**0.839**	
Corrugated iron sheets		470	150			
Other		568	186			
Construction material of walls	**1.243**	**0.91**	**1.697**	**1.873**	**0.171**	
Mud/Mud bricks		862	268			
Other		176	68			
Do you trust the community distributors	**1.105**	**0.841**	**1.452**	**0.515**	**0.473**	
I trust them a lot		757	239			
Other		278	97			
Do you trust people from different tribes	**0.825**	**0.625**	**1.089**	**1.847**	**0.174**	
I trust them a lot		250	93			
Other		792	243			
Do you trust people from other African countries	**1.182**	**0.833**	**1.677**	**0.879**	**0.348**	
I trust them a lot		168	47			
Other		874	289			
Household has access to mobile phone	**0.954**	**0.735**	**1.239**	**0.124**	**0.725**	
Yes		690	226			
No		352	110			

Table 5-22: Household level sociodemographic variables and their association with recent adherence to CDTI.

5.4.12 Variables independently associated with compliance to CDTI

The 12 variables that were identified as having a significant association with recent compliance were tested in a fully adjusted logistic regression model. From this model, I identified 6 variables that were independent predictors of compliance with CDTI. A subsequent step-wise model selection procedure using the 6 identified variables confirmed that the model with the lowest Aikake Information Criterion (AIC) was the model which included all 6 identified variables. The odds ratios, 95% confidence intervals likelihood ratio test χ^2 statistic and p-values are displayed in **Table 5-23**.

Variable	OR	L95%CI	U95%CI	LR Chisq	p-value	signif
Sex					**0.0220**	*
Female	0.58	0.36	0.93	5.25		
Male	1.00					
Currently Married					**0.0250**	*
No	0.59	0.37	0.94	5.03		
Yes	1.00					
Lived in village ≥ 10 years					**0.0005**	***
Yes	2.51	1.5095	4.1044	12.2012		
No	1.00					
Attended last village meeting					**0.0039**	**
No	0.45	0.2525	0.7805	8.3249		
Yes	1.00					
Household keeps animals					**0.0054**	**
No	0.5054	0.3181	0.8137	7.7571		
Yes	1.0					
Work on river/ocean					**<0.0001**	***
Yes	0.1255	0.0546	0.2922	21.1781		
No	1.00					

Table 5-23: Independent predictors of compliance included in a fully adjusted logistic regression model.

Females were less likely to be compliant than males. Being currently married was a positive predictor of compliance to CDTI, as was having lived in the village for 10 or more years. Persons who did not attend the last village meeting had reduced odds of being compliant as did households which did not keep animals. Finally, those that declared the river as their main place of work (fishermen), were nearly 8 times as likely not to have complied with CDTI in the last 12 months, all other factors being equal.

5.5 Conclusions

The surveys undertaken were used to identify factors that are independently associated with compliance to community-directed treatment with ivermectin and to find factors that are associated with self-reported onchocerciasis sufferers. Similar studies have been undertaken in APOC regions of Africa. In particular, concerning onchocerciasis, studies have been conducted in Ethiopa (*269*), Nigeria (*270, 271*), Cameroon (*270*) and Uganda (*267, 272*). Studies investigating factors associated with compliance for neglected tropical diseases other than onchocerciasis, both within and outside of the African setting, include studies on lymphatic filariasis (Haiti (*273*), Philippines (*274*), Kenya (*275*), India (*276, 277*)) and schistosomiasis (Kenya (*278*)).

To my knowledge these are the first data on predictors of compliance in the West African setting of the former Onchocerciasis Control Programme area, where vector control was established in 1975 and mass drug treatment with ivermectin began in 1987 (initially with mobile teams, before moving to the community-directed model). The method and duration of control in the study area is unlike any of the other studies cited (in APOC regions for instance, CDTI started in 1995 and there was never vector control except in Equatorial Guinea, Uganda and Tanzania).

In this study, I identified 6 independently associated predictors of compliance to community-directed treatment with ivermectin. I also identified several statistically significant associations between

various socio-economic factors and a participant positively identifying themselves as having experienced problems due to onchocerciasis.

One of the main weaknesses of the study is that firstly, I used self-reported compliance, which could be subject to recall bias, and secondly, I used short-term compliance as the outcome measure. It would have been better to verify each persons' actual compliance with CDTI by checking their record in the community distributor records. These were not always available, and not always complete. Another problem is that the records in the distributor notebooks did not match the figures reported at a national level. By way of example, in Jagbenbendo, where some records were available, comparing the summaries of CDTI for the previous 5 rounds of CDTI, the figures in the books were different from the aggregated figures provided at the national level by the Ghana Health Service. There were treatment rounds in the registers that did not appear in the national level data (see **Figure 5-17** where there were 3 treatment rounds in 2010). The overall treatment coverages also did not match national records (in this instance the community registers have a lower treatment coverage and indicate more treatment refusals). It could be possible that the third treatment round was an additional distribution to try to cover those that were missed during the July treatment. An additional problem could be that whilst distributors are usually literate, and are selected for that reason, sometimes they are not. The previous distributor in Nterso prior to the commencement of this study was illiterate, and was selected by the community. The present distributor is literate and was selected by health officials. Verifying compliance and coverage on the treatment registers should be a priority for any future study investigating coverage and compliance.

5 of the 6 independently identified factors were individual level variables, whilst the remaining factor was a household level variable. Gender was significant, with females having lower odds of complying with treatment. This was expected due to pregnant or breast-feeding women not being eligible for treatment and in a bivariate analysis of persons aged 16 and under the association is not significant (OR= 0.854 (0.626-1.165), p-value=0.319).

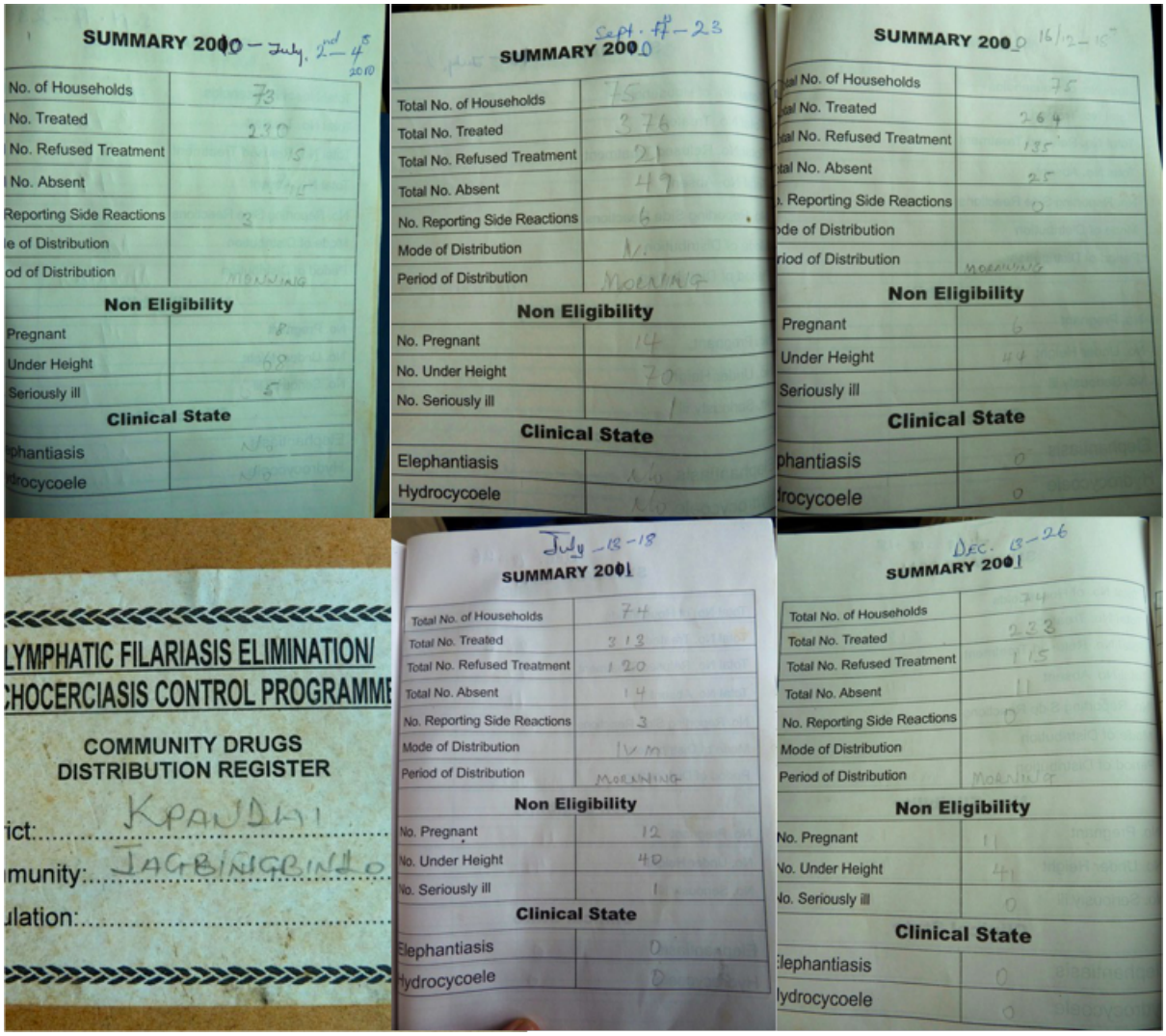


Figure 5-17: Summary pages from treatment registers at Jagbenbendo.
These were the registers that were being carried by the community distributor who was distributing ivermectin on the same day we were conducting the interviews. 3 treatment rounds are recorded for 2010, whilst only one treatment round for Jagbenbendo appears in national records for 2010 (see **Table 5-1**).

Being currently married was positively associated with compliance. Other compliance studies also found this association to be statistically significant (*270*) whilst others have not (*267*). Married couples may tend to be less mobile linked with raising a family, and therefore may spend less time away for work, and having a partner increases the chance that someone will be at the household when the distributor arrives in the house-to-house model of drug distribution, or who can serve as a reminder to collect their drugs in the central-site model of drug distribution (households go to a central point in the village where drugs are distributed and a record of the distribution is taken).

In this study, duration of stay within a village of greater than 10 years was a positive predictor of compliance. This same variable was however not significantly associated with the outcome measure 'Did you suffer onchocerciasis in the last 12 months?' (OR = 0.784, 95% CI=0.505-1.216, p-value=0.276). A possible explanation for this is that migration could be to and from areas where transmission is ongoing so more recent immigrants to a village may arrive already infected. There appears to be ample migration and resettlement within the studied region. Some of this may be seasonal, economic migration (people looking for work), but other migration (and particularly resettlement) is due to avoidance of disease. Asubende in the Brong-Ahafo region had the highest prevalence of blindness in the study: 4 times higher than anywhere else. However, Asubende has reduced dramatically in size over time.

During the 1990s, up until 2000 the population was relatively stable at around 550-600, however since then the population has plummeted to just 87 people in 18 households (see **Figure 5-18** for an outline of Asubende village at the time of the surveys). The team was unable to interview a full 20 households in this community, instead an additional 2 households were interviewed in nearby Senyase. During interviews with the remaining village Elders, they stated that many people had moved away from Asubende due to the blackfly problem, and associated blindness.

At least three new communities had been formed through migration out of Asubende; Fawoman, meaning 'I take my state' or 'I don't want anything to do with you'; Matukwa, meaning 'I left with a reason' and Mempesem, meaning 'I don't want any trouble'). The names are references to onchocerciasis. The people remaining in Asubende may be those who are most vested in the village (e.g. the village Chief) and those who are already permanently disabled through exposure to biting blackfly and *O. volvulus*.

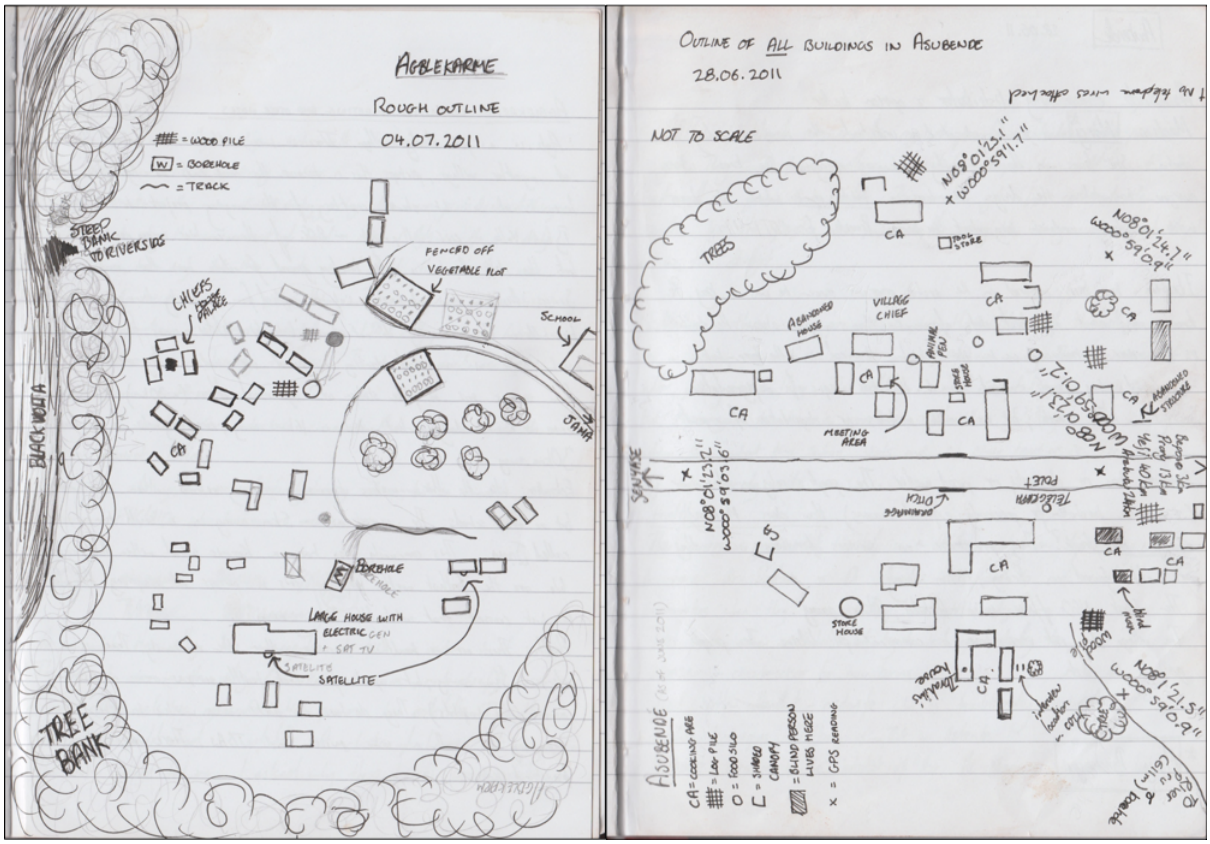


Figure 5-18: My outline sketches of the villages of Agblekeme and Asubende.
Sketches show some of the typical features of each village and the location of the rivers in location to housing. The map of Asubende depicts the location of all inhabited houses in the village at the time of the survey (28th June 2011).

The variable 'Attended last village meeting' was a positive predictor of compliance in this study. This could be a proxy for social engagement and integration in a community. Some of the study sites are spread out over large areas, and it may be difficult for a distributor to reach some of those houses, particularly for distributors who cover a large area and multiple villages.

In Agblekeme a family who were interviewed stated that had not taken ivermectin because the distributor never visits their house. Their compound was sited some way outside the main settlement, approximately a 10 minute walk from the centre of the village. There was no effect of other measures of social integration on an individual's compliance with CDTI. There was no association with a household's opinion of the community distributor (see **Figure 5-11**) and their willingness to comply with treatment.

Individuals from households who kept animals were about twice as likely to comply with CDTI as those who did not have animals. This could be because those with animals need to stay in and around the village to effect day-to-day husbandry tasks of livestock. Many families are subsistence farmers and the livestock they own and raise are an important source of food and marketable goods and therefore they are unlikely to be left unattended in communities.

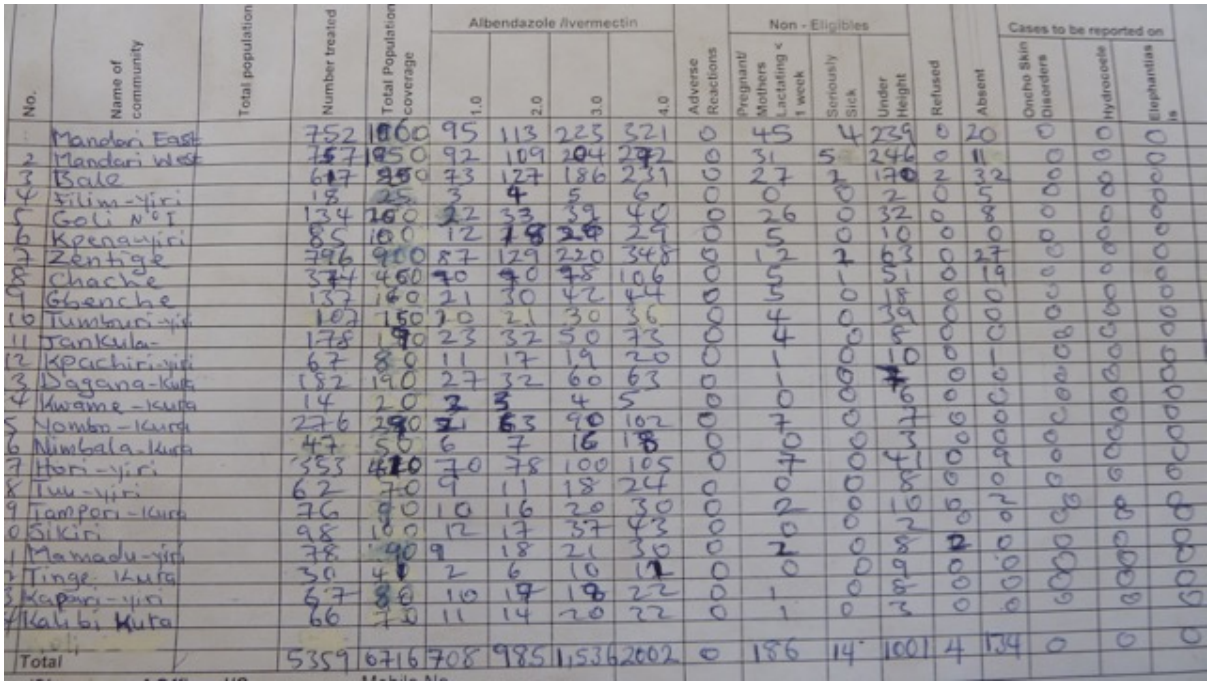
Finally, the independent predictor 'working on the river' as one's main place of work was highly associated with a lack of compliance. Individuals who worked on the river were nearly 8 times as likely not to have complied with CDTI last year than those who did not work on the river. Those who work in the fishing industry may spend extended periods of time away from their households on fishing trips, and they may also participate in seasonal economic migration, moving away from the village for several days or weeks at a time in order to undertake a fishing trip.

In the analysis of different types of compliance, several villages were found to contain so-defined systematic non-responders. Two of these villages were previously identified in Osei-Atweneboana et al (*113*) as being 'of concern' as villages with infected hosts who had a suboptimal response to ivermectin (microfilarial repopulation rates were greater than expected). This finding was confirmed by Churcher *et al*. (*110*) who found a statistically significantly faster rates of skin repopulation in Jagbenbendo, Kyingakrom and New Longoro. The suboptimal responses identified in Kyingakrom and New Longoro could partially be explained by the systematic non-compliance found in these villages in this study. Between 6 and 6.5% of persons in these villages were found to be systematic non-compliers in this study. Chache had an unusually high proportion of systematic non-compliers. Many households in this community had just been re-housed there, in new purpose built concrete dwellings, due to development at a nearby dam site in Bui National Park. This community was also one of two which did not have a hand in choosing their community distributor. Nsunua was the other community which did not choose their distributor, which also had a high rate of non-compliers. In the community leader's interviews at Senyase the elders indicted they did not elect the distributor, but the distributor lived in the next village (Asubende) and had been working as the distributor in these

villages since the onset of ivermectin distribution in 1987. Empowering the community to elect their own distributors appears to be crucial to ensuring a high rate of compliance with CDTI.

In Ghana, there are both forest and savanna forms of onchocerciasis, the former being associated with high prevalence of skin lesions, but a very low prevalence of blindness, whilst the latter is associated with high prevalence of blindness. The Northern and Brong-Ahafo regions of Ghana are both located within the savanna zone of Ghana and therefore it was not possible to write anything about the effect of bioclime on both prevalence of symptoms and compliance with CDTI in these areas.

Future research must align rich socioeconomic survey data with the raw community distributor treatment records to get an accurate count of the number of rounds of treatment an individual has received. An interesting study would be to audit a selection of community distributor records and follow the data as they are aggregated at district and regional level to ascertain where any inaccuracies in reporting may occur. An example of district level aggregation is shown in **Figure 5-19**. This was the record which was currently held at the Wiae Community-Health and Planning Service compound, which covers the Mandari sub-district of the Bole District in the Northern Region, was to be sent to the District level. All forms are initially filled out by hand before being entered onto a computer.



No.	Name of community	Total population	Number treated	Total Population coverage	Albendazole /Ivermectin 1.0	2.0	3.0	4.0	Adverse Reactions	Non - Eligibles: Pregnant/ Mothers Lactating < 1 week	Seriously Sick	Under Height	Refused	Absent	Cases to be reported on: Oncho Skin Disorders	Hydrocoele	Elephantiasis
1	Mandari East		752	1060	95	113	225	321	0	45	4	239	0	20	0	0	0
2	Mandari West		757	1850	92	109	204	272	0	31	5	246	0	11	0	0	0
3	Bale		617	950	73	127	186	231	0	27	2	170	2	32	0	0	0
4	Filim-Yiri		18	25	3	4	5	6	0	0	0	2	0	5	0	0	0
5	Goli N°I		134	160	22	33	39	40	0	26	0	32	0	8	0	0	0
6	Kpena-yiri		85	100	12	18	24	29	0	5	0	10	0	0	0	0	0
7	Zentige		796	900	87	129	220	348	0	12	2	63	0	27	0	0	0
8	Chache		374	460	70	90	78	106	0	5	1	51	0	19	0	0	0
9	Gbenche		137	160	21	30	42	44	0	5	0	18	0	0	0	0	0
10	Tumburi-yiri		107	150	20	21	30	36	0	4	0	39	0	0	0	0	0
11	Jankula-		178	190	23	32	50	73	0	4	0	5	0	0	0	0	0
12	Kpachiri-yiri		67	80	11	17	19	20	0	1	0	10	0	1	0	0	0
13	Dagana-Kura		182	190	27	32	60	63	0	1	0	7	0	0	0	0	0
14	Kwame-Kura		14	20	2	3	4	5	0	0	0	6	0	0	0	0	0
15	Nombo-Kura		276	290	71	63	90	102	0	7	0	7	0	0	0	0	0
16	Nimbala-Kura		47	50	6	7	16	18	0	0	0	3	0	0	0	0	0
17	Hori-yiri		353	420	70	78	100	105	0	7	0	41	0	9	0	0	0
18	Tuu-yiri		62	70	9	11	18	24	0	0	0	8	0	0	0	0	0
19	Tampori-Kura		76	90	10	16	20	30	0	2	0	10	0	2	0	0	0
20	Sikiri		98	100	12	17	37	43	0	0	0	2	0	0	0	0	0
21	Mamadu-yiri		78	90	9	18	21	30	0	2	0	8	2	0	0	0	0
22	Tinge Kura		30	40	2	6	10	12	0	0	0	9	0	0	0	0	0
23	Kapori-yiri		67	80	10	19	18	22	0	1	0	8	0	0	0	0	0
24	Kalibi Kura		66	70	11	14	20	22	0	1	0	3	0	0	0	0	0
	Total		5359	6716	708	985	1536	2002	0	186	14	1001	4	134	0	0	0

Figure 5-19: District level coverage record for Mandari sub-district, Bole district, Northern region.

5.6 Acknowledgements

I would particularly like to thank Dr Isaac Osei-Akoto at the Institute for Statistical, Social and Economic Research (ISSER) who provided logistics, support, guidance and field teams for undertaking this fieldwork. I would also like to thank George Domfe of ISSER who was the interview team leader and Paul Bego (ISSER), George Asare (ISSER), Evans Adomako Appiah, Raphael Agyare Nana and Maxwell (ISSER) who conducted the interviews in the local languages.

Chapter 6. General Discussion

The objective of this chapter is to a) frame the work presented within this thesis within the wider context of on-going efforts to eliminate onchocerciasis, b) to summarise the main findings of this thesis and c) to briefly outline future research directions that could build upon the results from this thesis.

6.1 Summary of work in this thesis

In the context of elimination of onchocerciasis, the major aims of this thesis were to map the pre-control distribution and prevalence of human onchocerciasis in West Africa, and to identify factors linked to coverage and compliance with community-directed treatment with ivermectin (CDTI). To generate robust maps of disease distribution, I carried out a detailed examination of spatial autocorrelation present in the OCP epidemiological dataset, and which methods may be best to map prevalence using this data (Chapter 2). From this investigation, I applied a Bayesian geostatistical methodology, also identifying important environmental covariates that could help estimating prevalence in areas distant from ground-truth survey data, generating both point estimates of prevalence and credible intervals around those estimates (Chapter 3). The pre-control level of endemicity is a good indicator of the feasibility of elimination using community-directed treatment with ivermectin (CDTI) alone. The mapped pre-control prevalence was combined with a population count map to estimate the population infected and at-risk prior to control in the study area, and estimates of the burden of onchocercal disease were made, using regression models to link prevalence of infection to prevalence of morbidity (Chapter 4). Finally, I carried out a hierarchical socioeconomic survey of 280 households in 13 villages in Ghana, West Africa, to find factors associated with coverage and compliance with CDTI (Chapter 5).

6.2 From control to elimination of NTDs

The NTDs are a growing list of predominantly chronic diseases that disproportionately affect the world poorest and most vulnerable people. Unlike many of the other major communicable diseases, left untreated NTDs result in chronic long-term disability, with increased incidence of death due to infection only occurring after many years (with exceptions such as visceral leishmaniasis), and lead to long-term disability, trapping those affected in a vicious cycle of chronic infection and on-going poverty. Several NTDs such as onchocerciasis have been the subject of successful, long-running control programs to control disease as a public health problem. However, to prevent recrudescence of morbidity, control must be either be maintained indefinitely, or control programs must focus efforts on elimination, establishing what the relevant transmission dynamic, epidemiologic and programmatic parameters are, to reach elimination endpoints. The establishment of the United Nations Millennium Development Goals (MDGs) provided the impetus for governments and non-governmental organisations (NGOs) to ring-fence funding to pursue elimination (rather than control) of a wide variety of NTDs, whilst later initiatives such as the 2012 London Declaration on NTDs, have strengthened this resolve and ensured the continued engagement between stakeholders (such as funders) and national governments. There now exists both a scientific and moral obligation to ensure that these programmes can deliver on the promise to eliminate these communicable afflictions of poverty.

6.3 Model-based Bayesian geostatistical mapping

6.3.1 Summary of mapping pre-control prevalence of onchocerciasis in West Africa and main findings of study

In this study, I used parasitological prevalence data collected by the OCP prior to the commencement of control operations, and environmental covariates to estimate disease prevalence in un-sampled

areas, without ground-truth data (i.e. where no surveys took place). Though the OCP surveyed extensively within the control program's boundaries, covering a large geographic area, parasitological examination is extremely time-consuming, expensive, not sensitive at low infection intensity, and - due to the invasive nature of the skin snip method - not desirable by those being surveyed. Due to these factors, only a small fraction of the total villages and settlements were ever surveyed. Comparing Figure 2-1 in Chapter 2 showing the location of all surveyed locations in the OCP epidemiological dataset, with Figure B-7 in Appendix B showing the location of all settlements in just the Upper Volta region of Ghana, there are many, many more settlements in the Upper Volta region alone than those that could be surveyed in that area. Spatial autocorrelation present in the parasitological data (surveys points closer together tend to be more similar than those farther apart) meant that I could use Bayesian model-based geostatistical techniques to make estimates about prevalence. My model achieved a correlation coefficient of 0.69 between estimated and observed values in a cross-validation exercise, with a mean absolute prediction error of 12%. I estimated that in the OCP program boundary, 7.55 million people were infected prior to the start of control, with an average prevalence of 45%. In addition, I found distance to nearest river source to be the most informative predictor in the model, which is very relevant in the lifecycle and transmission dynamics of human onchocerciasis. Several other environmental covariates, which can be linked to seasonality and the likely persistence of vector populations were also informative. Pre-control transmission 'hotspots' were identified through this study. These are areas that are highly hyper-endemic for onchocerciasis. In Ghana, overlaying data from the last available REMO surveys (Figure 5-2), it was clear that these pre-control transmission hotspots are linked to ongoing transmission, even after 25+ years of treatment with ivermectin.

6.3.2 Limitations of the study

Two of the major issues I faced during this study were a) the sampling design of the study data itself and b) the resolution of the predictive surface and the available environmental explanatory covariates. The initial aim of the OCP was to attempt to find all areas where the disease was present, using

prudent expert knowledge about the entomological conditions prevalent throughout the program area to preferentially sample villages were onchocerciasis disease was likely present. Because the vectors breed in fast-flowing streams and rivers, initial surveys were carried out in all those villages classified as first-line villages which were those closest to river streams (see Section 3.5.1). Using distance to river as a covariate in the geostatistical model help to account for the non-random sampling design of the survey data

Other riverine factors were not found to aid the predictive power of the map – namely using Strahler stream order (a measure of the magnitude of a stream segment) and whether a stream segment was an intermittent or a perennial water body. It is possible that the variation in prevalence explained by these factors may have been better explained by other environmental covariates linked to precipitation and temperature, such as mean temperature of the wettest quarter and precipitation seasonality. However, it should be noted that almost all remotely sensed data products are themselves the result of models based on the input data (e.g. interpolation between weather station data such as used in BIOCLIM (*191*)), with associated error, modeling assumptions and finite resolution, and are therefore an imperfect measure of the true value at sampled points of interest.

This study was particularly challenging because remotely sensed data needed to be applicable to 1975 for this study. The earliest satellite data products started appearing in the late 1970s (such as the GLIMMS data products for normalized difference vegetation index [NDVI]), but these had a relatively low resolution (8km x 8km cells, or 64Km^2). In this study, I made the decision to use more up-to-date measurements, made at a finer resolution despite them not being sampled during the year of interest. I made the pragmatic choice that by averaging across many years of data I would be using the long-term average for indices such as NDVI which may be more applicable for modelling the pre-control endemic equilibrium, rather than a single year of data at low resolution using older equipment. Altitude and distance to nearest stream segment are also data products based on models (though the data layer for altitude is available at extremely fine spatial resolutions). These layers are more invariant with respect to short-term changes across time.

My output maps were produced at 5x5km resolution (25km^2). Mapping (including sampling the values of environmental covariates) at this resolution may have hidden some of the finer-scale spatial heterogeneities present in the relationship between infection prevalence and environmental covariates. The increasing volume and greater access to large compute resources (many thousands of CPU cores) in high performance clusters (HPCs) should make mapping at finer spatial resolutions more tractable for future iterations. Mapping at finer spatial scales also requires the concomitant production of higher resolution spatial data. One example of this is the on-going use of machine learning techniques applied to ultra-high resolution (0.5m) satellite imagery to produce very high resolution maps of population aggregation (**Figure 6-1**). These kinds of maps will enable much better estimates of the population infected and at risk for diseases which are transmitted within environmental niches (such as onchocerciasis, where transmission intensity is highest in those populations that live closest to breeding sites).

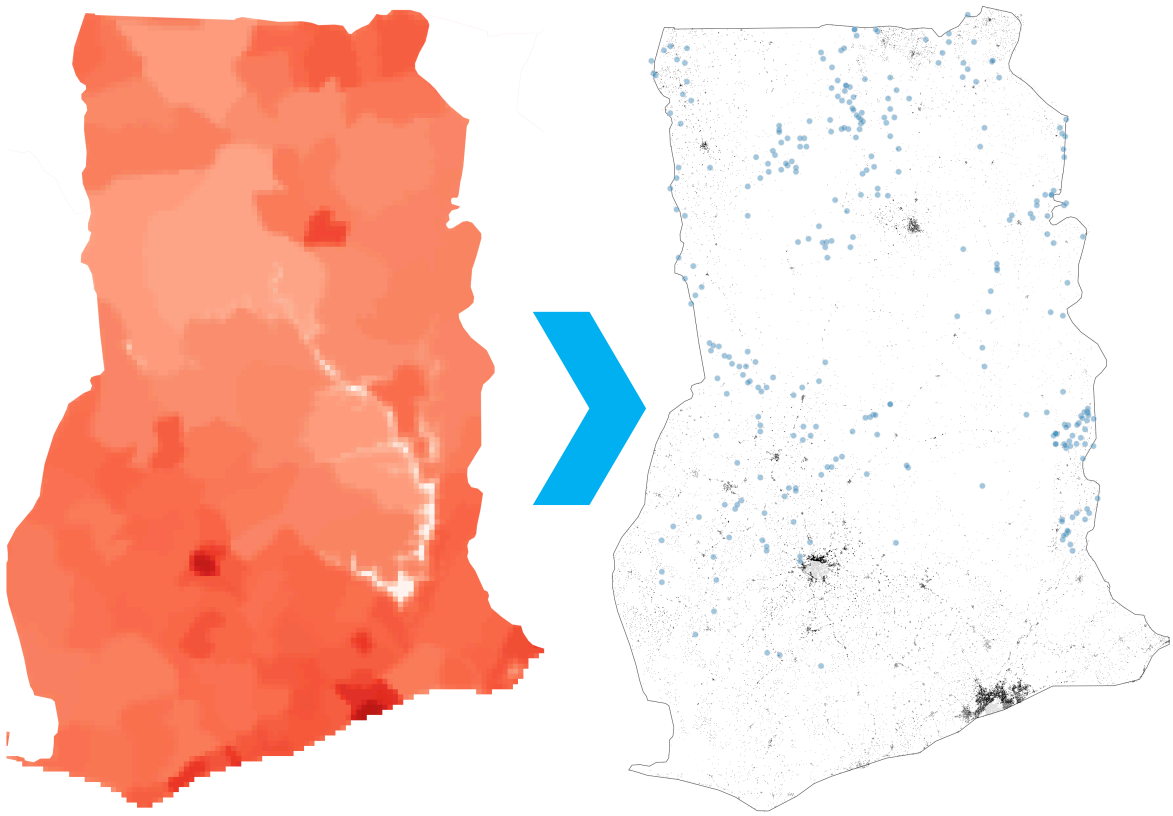


Figure 6-1: Recent updates in population distribution maps.
Machine learning techniques are being applied to ever more high resolution satellite imagery to produce maps of population aggregation in different global settings. The map on the left is the gridded population count as supplied by the third iteration of the Gridded Population of the World project (GPWv3) at the 25km2 resolution used as input to the maps in Chapter 3. The map on the right uses machine learning to identify individual structures in 0.5m resolution imagery, distributing count data from censuses into the structures identified, producing much more accurate maps of population aggregation. The blue dots overlaid on the map on the right are the locations of all villages that were surveyed during the OCP in Ghana.

6.3.3 Further uses of these maps to inform NTD elimination programs

The maps presented in Chapter 3 represent are first continuous maps of the pre-control prevalence of *O. volvulus* microfilariae in West Africa. Mapping baseline endemicity provides an indicator of initial conditions for transmission, which directly affect the feasibility of onchocerciasis elimination with ivermectin (*109, 140*). In the context of current efforts to eliminate onchocerciasis, these pre-control maps have not only historical value, but are a very pertinent tool to better target current interventions, and to guide where increased monitoring and evaluation may be necessary to detect and prevent recrudescence (*279*) once elimination is achieved.

Maps including the ones I generated are being used in current research led by the NTD Modelling Consortium (NTD MC) and the London Centre for Neglected Tropical Disease Research (LCNTDR),

to integrate mapping with dynamic disease transmission models (http://www.londonntd.org/projections-on-eliminating-ntds-integrating-mapping-with-modelling), projecting forwards in time, incorporating knowledge about interventions, to make predictions about when disease might be eliminated, and what level of intervention would be needed to achieve those goals.

There is also an increasing effort to integrate control programmes to lower costs, improve logistical efficiency and to accelerate the progress towards elimination for multiple diseases through the synergistic action of integrated drug regimens (*280, 281*). However, a barrier to this for onchocerciasis is the risk of severe adverse events in persons co-infected with loiasis (*282*). Co-endemicity maps are therefore required to inform national control programs about which regions are safe to treat with control programs that include ivermectin. Since publication, my maps have been integrated with other maps of onchocerciasis (*216*), loiasis (*138*) and lymphatic filariasis risk (*283*) to develop filariases co-endemicity maps for the whole of sub-Saharan Africa (SSA) (*284*).

My maps are also publically available on www.ntdmap.org, an interactive mapping tool (**Figure 6-2**). The purpose of ntdmap.org is to enable decision makers and other users, who may have little geographic information systems (GIS) knowledge to produce or consult disease maps to inform programmatic decision making (*285*).

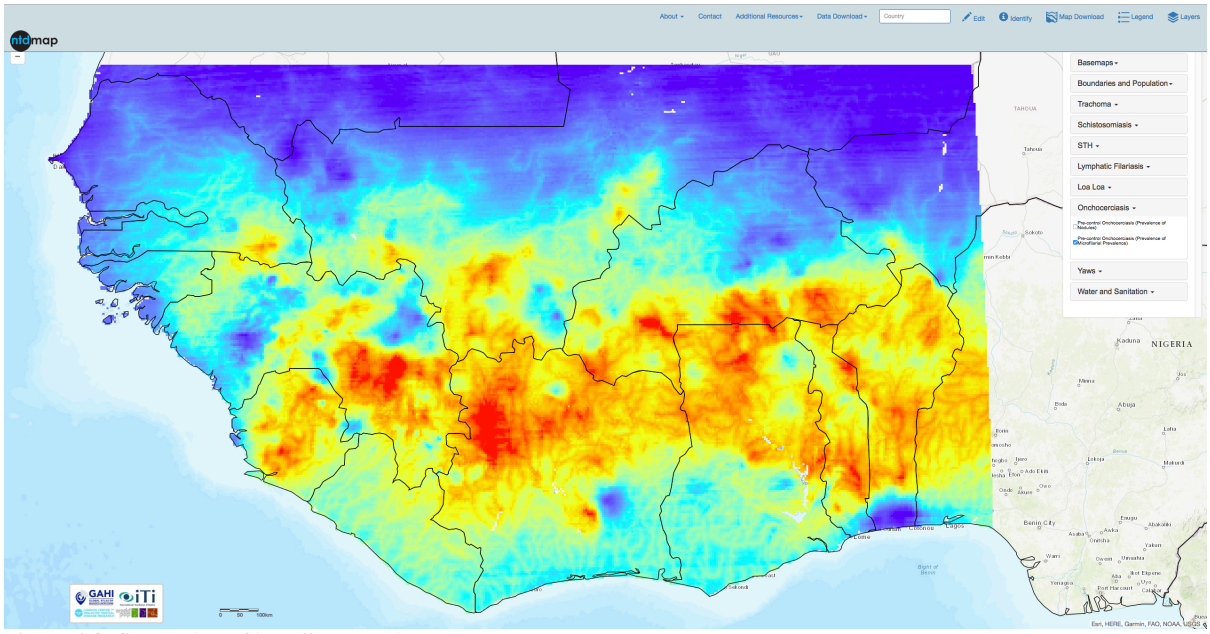


Figure 6-2: Screenshot of http://www.ntdmap.org.
Displayed is my mean microfilarial prevalence of *O.volvulus* in the population aged 5 and above from the Bayesian geostatistical modelling carried out in Chapter 3. This is displayed as a data layer, which may be overlaid with other data layers as necessary.

6.3.4 Other mapping approaches to support disease control and elimination programs

Increasingly, maps other than those of disease prevalence are being used in planning for disease control and elimination. Lasting solutions to global healthcare problems cannot be solved without significant concurrent investment in infrastructure, and particularly access to clean water, sanitation and hygiene (WASH) facilities and greater access to healthcare for those in the most rural and disadvantaged populations. Mapping of drinking water supply and sanitation facilities helps identify those populations where further investment in those infrastructures may be necessary to help control infectious diseases and has been carried out for Sub-Saharan Africa (*286*). Functional water sources in Ghana were mapped using spatial methodology to identify communities in need of improved water access, particularly in this case to combat schistosomiasis risk (*287*). Human migration affects the risk of recrudescence of disease and disease importation into areas where local elimination has been achieved. Migration is particularly relevant in the setting of sub-Saharan Africa (SSA) where intraregional migration is common (*288*). The huge increase in mobile phone ownership in SSA over the past decade has created a new source of digital data which has been leveraged to track human movement patterns, at very fine spatiotemporal scales, unconstrained by the logistics of carrying out

national censuses (*289*). It is now possible to look at human movement over short time-scales (e.g. day/night) and longer time scales (e.g. weekday/weekend, school-term/holiday, summer/winter) and use these to inform multiple control program issues such as targeting interventions (*290*), health impact assessments (*291*)and mathematical modelling of transmission dynamics(*292, 293*). The integration of multiple sources of data will better enable the integration of geospatial and mathematical modelling techniques to inform policy and control program planning in the future.

6.4 Disease Burden Estimation

6.4.1 Summary of main findings

I used the mean prevalence of mf infection map calculated in Chapter 3 as a numerator to make an estimate of the burden of disease caused by onchocerciasis at the onset of the OCP. The denominator had to be estimated from gridded population maps which present per-pixel population estimates, but for the year 2000. To make the denominator relevant for this study I needed to apply a back-transformation. I transformed the population count in each pixel using country and time-specific population growth rates. Summing the transformed population pixels for each country I found less than a 0.8% difference between my estimated pixels and the United Nations 1975 census estimates (which were only available at country level).

I estimated 7.55 million people were infected with onchocerciasis (within the OCP program boundaries) at the onset of the OCP, with 17.8 million people 'at-risk' of onchocerciasis in these areas (the total estimated rural population at the time). I also estimated that there could have been an additional 11.6 million people at-risk of onchocerciasis in OCP countries, but living outside the program boundaries, of which 2.8 million may have been infected (using the mean prevalence map from Chapter 3).

I made an estimate of the burden of onchocercal skin disease using the estimated numbers of persons infected. Using models which link the prevalence of infection to prevalence of morbidity of skin disease I estimated that in the OCP program boundary, around 350,000 DALYs were lost due to onchocercal skin disease in 1975 at the inception of the OCP. The GBD 2010 study estimated a global total of 490,000 DALYs lost to onchocerciasis (with 100% of this total due to years living with disease, rather than any years of life lost) which includes the effects of 35 years of population growth.

6.4.2 Limitations of disease burden estimation and extensions to the study

In this study, I made estimates of the years living with onchocercal skin disease. An obvious extension to this work would be to include the sequelae of vision loss and blindness, which would comprise a substantial proportion of the total DALYs lost due to infection with onchocerciasis prior to the start of intervention programs. There are other sequelae which could also be included in burden estimates, including onchocercal epilepsy (*66, 294, 295*) and excess mortality linked to intensity of infection (*4*). These additional sequelae will contribute a large proportion of the total DALYs lost due to onchocerciasis and should be included.

A drawback of this study is the use of prevalence of infection to estimate morbidity. A better marker of morbidity (in the pre-control endemic equilibrium setting) would be the use of intensity. Whilst there exists a linear relationship between the intensity of infection in the community and the prevalence of sequelae of ocular onchocerciasis, the relationship between eye disease and prevalence is strongly non-linear (*25*). Most ocular disease occurs in hyperendemic communities (with a mf prevalence of 60% and above), however mf prevalence may not well explain differences in prevalence of eye disease between hyperendemic communities, due to large variations in the intensity of infection at high prevalence (**Figure 6-3**). As such, any estimates of vision loss based on prevalence of infection will incorporate greater uncertainty than estimates based on intensity.

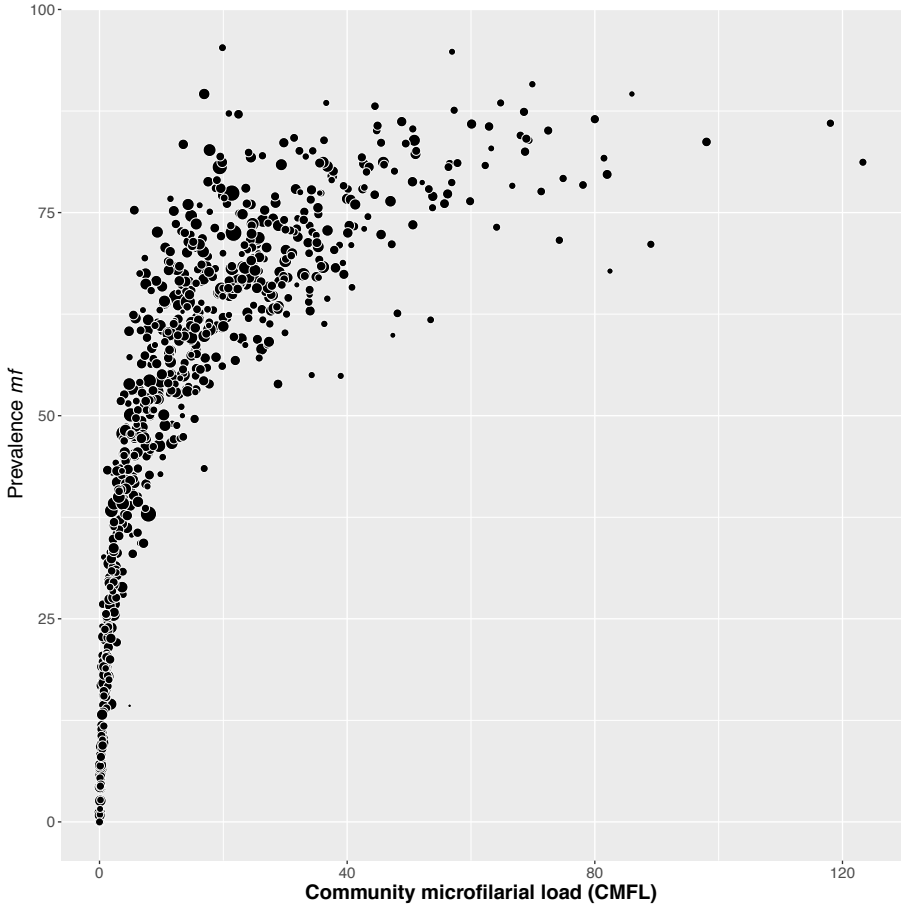


Figure 6-3: Variation in the relationship between prevalence and intensity of infection with mf.
Each point represents a village and are the pre-control data from the OCP epidemiological dataset. Points are scaled by numbers of people examined in each village.

A further extension to pre-control burden estimates incorporating all disabling sequelae, would be to carry out a counterfactual analysis. In one scenario, a model would project forwards in time, incorporating the effects of population growth to estimate the disease burden in the absence of the efforts of the OCP. The assumption is that the endemic equilibrium is stable, and that disease burden will be driven solely by fluctuations in the population and the rate of migration of people to and from onchocerciasis endemic and non-endemic areas.

The alternative scenario is to incorporate the effects of the OCP to investigate the reduction in disease burden prevented by the intervention program. In addition to demographic factors, modelling efforts in this setting would need to account for the duration and type of intervention implemented in each

area. Larviciding of vector breeding sites, which was widely implemented in the OCP between 1974 and 1987, has a slow reduction on microfilarial density in the skin of infected persons (there is a large reduction in infective bites so no new adult worms can establish – the reduction in microfilarial density occurs through mortality of existing adult worms). The introduction of ivermectin in 1987 quickly led to a rapid reduction in intensity of infection (ivermectin is a highly efficacious microfilaricide, but as adult worms are not killed by ivermectin, high coverage is required over many years to interrupt transmission long enough to achieve elimination). Models therefore need to account not only for the initial prevalence of the different types of morbidity, but also what the effect of the different types of interventions would be on both existing disability and the incidence of new cases of disabling sequela. The incidence of new cases non-reversible disease which is caused by long-term cumulative exposure to mf would be greatly reduced (but not stopped) by the larviciding of vector breeding sites. Individuals who are yet to suffer the latter stages of non-reversible morbidity, but whom already have a high microfilarial load, may go on to develop severe vision loss, blindness and severe skin disease because their cumulative exposure to mf would not be reduced until the introduction of ivermectin in 1987. The use of ivermectin from 1987 onwards leads to an almost complete reduction in the incidence of new cases of skin disease, vision loss and blindness due to onchocerciasis (because afflicted persons instant and future exposure to mf is almost completely prevented). Ivermectin will also cause a rapid reduction in the prevalence of less severe skin disease such as troublesome itch. However systematic non-compliers may help to harbor parasite reservoirs in endemic communities, and since vector control was almost completely stopped after the introduction of ivermectin, onward transmission of new parasites by infective female blackflies would be possible.

Although community-directed treatment with ivermectin has had a remarkable impact on the incidence of new cases of onchocercal skin disease, it is clear from recent parasitological surveys (*262*) that transmission is an ongoing problem in some part of Ghana. The socioeconomic surveys that I undertook in Chapter 5 have shown that there are still significant problems with onchocerciasis morbidity, with 22% of people in the study region registering complaints about onchocerciasis morbidity, and 1.8% of persons blind.

6.5 Coverage and compliance with community-directed treatment with ivermectin

6.5.1 Summary and main findings of the socioeconomic surveys carried out in Ghana, West Africa

Establishing which individual, household or community level factors may be associated with poor coverage can help inform program managers in national control programs about best-practices for designing control programmes to maintain high-levels of therapeutic drug coverage and long-term compliance. To investigate reports of suboptimal responses (SORs) to community-directed treatment with ivermectin (CDTI), I collected data on the socioeconomic setting of 1383 individuals in 280 households across 13 villages in 2 regions of Ghana, West Africa. I collected information on their onchocerciasis status and their long-term history and recent compliance to treatment with ivermectin.

I found that 5 individual and 1 household level variables were independently associated with increased compliance to treatment with ivermectin: gender (being male); being currently married; duration of stay of greater than 10 years in the village; having attended the last village meeting; a river or ocean not being one's main place of work and finally, that the household does not keep animals. I also found that systematic non-compliance could be highly variable between villages, ranging from 0% to 25%.

Model predictions about the expected impact and efficacy of interventions are sensitive to assumptions surrounding drug compliance. Systematic non-compliers (those who never take ivermectin) can have a huge impact on the feasibility of elimination, by creating a parasite reservoir in the community from which transmission may be on-going (there is almost no vector control to guard against the biting of diurnal *Simulium* spp.). A widely-used assumption for the level of systematic non-compliance used in transmission dynamic modelling of onchocerciasis is to assume an upper-bound of 5% of persons are systematic non-compliers (*108, 109, 184, 279, 296*). In the data I

collected, I found that systematic non-compliance could be as high as 15-25%, in villages which did not have a stake in choosing their community distributor (sometimes distributors are appointed by the health ministry covering multiple communities). This markedly affects the chances of elimination for a given transmission setting, even after 30 rounds of CDTI at high coverage (**Figure 6-4**).

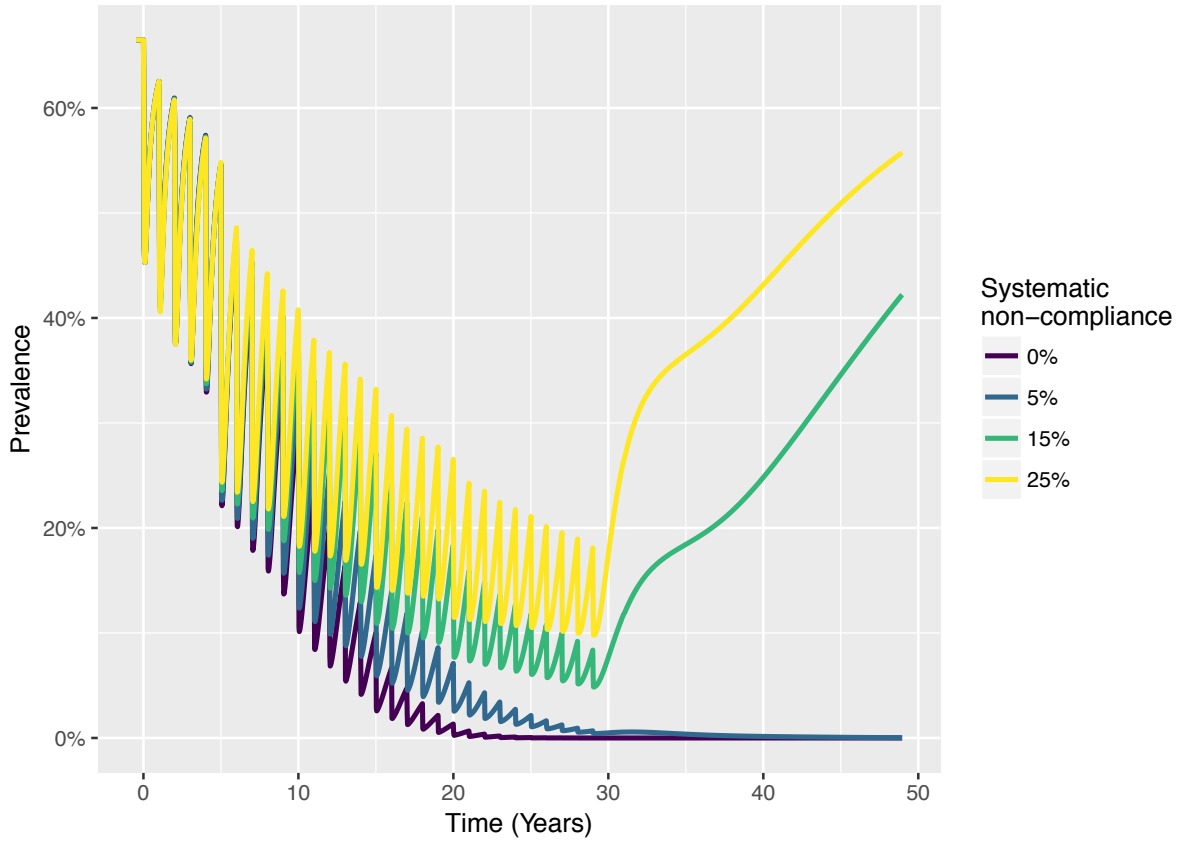


Figure 6-4: Simulation of 30 rounds of CDTI using EPIONCHOv2 (*279*) under different assumptions of systematic non-compliance.
Under the assumption of 0% or 5% of persons systematically refusing treatment, elimination can be achieved after 30 rounds of CDTI. If the proportion of systematic non-compliance is higher, microfilarial prevalence will rebound to pre-control levels, even after 30 rounds of treatment (therapeutic coverage is the same in all scenarios and is assumed to rise from an initial value of 50%, up to 80% during the first 20 annual rounds of CDTI and stays fixed at 80% for the final 10 rounds of CDTI). Other transmission parameters were as described in (*279*).

Several sub-optimally responding (SOR) villages in response to community directed treatment with ivermectin were identified by the longitudinal study of Osei-Atweneboana et al (*113*) and three of these were confirmed to have statistically significant faster than expected repopulation rates (*110*). However, in the socioeconomic study that I undertook in the Northern and Brong Ahafo regions, I visited these SOR communities in addition to twice the number of normally responding communities. In two of the three identified SOR communities I found that there was cause for concern over

peoples' attitude towards CDTI and their adherence to treatment. Kyingakrom and New Longoro which were villages of concern, both had 6% systematic non-compliance. In addition, I found that treatment coverage was highly variable, and that data obtained directly from the national control program, conflicted with the raw data I could see in village treatment records. Models predicting the effectiveness of CDTI may need to be revised to incorporate these findings and additional uncertainty in coverage treatment records.

6.5.2 Limitations of the study

There were three major limitations of this study. Firstly, all the data collected were self-reported by interview participants and were not validated against community distributor records in any way. I collected information on compliance by asking about a person's compliance with the last round of CDTI. This was assumed to be less subject to recall error than asking for someone's complete history. However, I designated someone as a random or systematic complier based on their reason for missing the previous treatment round and their reasons or attitude to compliance may change over time.

The second major limitation of this study is whether the savannah-type onchocercal bioclime of Ghana (where this data were collected) is applicable more broadly to other onchocerciasis endemic settings, for instance in East and Central Africa and in regions where the forest-type onchocerciasis is prevalent. Other limitations include selection bias – household were selected randomly for the study from a complete household listing of each village. However, some households may refrain from participating for several reasons which may lead to self-selection bias by households that choose to participate in the study. Overall, since only very small number of households refused (and this was due to economic commitments or planned absences on the survey day).

The other main limitation in this type of study is that we rely on cross-sectional observation to examine long-term trends, i.e. an individual's willingness to participate in treatment programs. Individual preferences and willingness to participate may change across time, especially in the context

of infectious diseases where individuals may only display treatment seeking behavior once they display symptoms of disease. Of the 336 individuals who did not take ivermectin during the last treatment round, 17 (5%) of those individuals claimed it was because the drug was not useful. A longitudinal study may help to better examine long-term trends and willingness to participate across time. In the study design a major question would be to use a cohort-based follow-up study, or to conduct cross-sectional observational studies over multiple years with an opportunistic cohort. An issue with a follow-up based study is that the persons in the study could be biased to take ivermectin due to inclusion in the study.

6.6 Future research directions

With the focus of control efforts aimed at eliminating infection with onchocerciasis, a key question will be, whether it is possible to eliminate onchocerciasis with annual or bi-annual ivermectin distribution. The pre-control distribution of microfilarial prevalence maps can inform program managers where highly hyperendemic foci existed prior to interventions that may require the implementation of special intervention zones (where limited vector larviciding or increased drug treatment targeting may also be required). In mathematical models of the feasibility of elimination with current tools, a key assumption is the level of compliance with CDTI. My thesis found that this can be highly variable, ranging from 0-25% systematic non-compliance in different communities which greatly affects the probability of elimination even after many rounds of ivermectin treatment at high assumed coverages.

Future research directions could include the use of new and sophisticated remotely-sensed data to inform geostatistical models of disease distribution, at very fine spatial resolutions. Dynamic models that integrate seasonality in environmental covariates, and human population dynamics (migration and demographic changes) could inform efforts to update pre-control maps taking account of control programme interventions. Additional studies (including longitudinal studies) on coverage and

compliance to CDTI, could help to inform mathematical models about long term trends of systematic non-compliance in endemic regions and what socioeconomic risk factors drive low adherence to treatment programs.

Despite the considerable progress made towards elimination of onchocerciasis in Africa under the auspices of the OCP and APOC (and subsequently national control programmes), there remains ongoing transmission which may prevent the eventual elimination of onchocerciasis if not addressed. To achieve elimination and to avoid recrudescence of this once widely-disabling disease, further research will be required into what diagnostic measurements can safely be used to confirm true elimination in widespread transmission settings, without risk of recrudescence from other areas with infection prevalence.

References

1. London Declaration on Neglected Tropical Diseases, "Uniting to combat neglected tropical diseases: Ending the neglect and reaching 2020 goals.," (2012).

2. M. E. Murdoch *et al*., A clinical classification and grading system of the cutaneous changes in onchocerciasis. *Br J Dermatol* **129**, 260-269 (1993).

3. J. Bradley, Whitworth J, B. MG, in *Topley & Wilson's microbiology and microbial infections,* F. E. G. Cox, D. Wakelin, S. Gillespie, D. D. Despommier, Eds. (Arnold, London ; New York, 2005), vol. Parasitology, pp. 781-801.

4. M. P. Little, L. P. Breitling, M. G. Basanez, E. S. Alley, B. A. Boatin, Association between microfilarial load and excess mortality in onchocerciasis: an epidemiological study. *Lancet* **363**, 1514-1521 (2004).

5. M. Walker *et al*., Density-dependent mortality of the human host in onchocerciasis: relationships between microfilarial load and excess mortality. *PLoS Negl Trop Dis* **6**, e1578 (2012).

6. A. P. Plaisier, G. J. van Oortmarssen, J. H. Remme, J. D. Habbema, The reproductive lifespan of Onchocerca volvulus in West African savanna. *Acta Trop* **48**, 271-284 (1991).

7. World Health Organization: Expert Committee on Onchocerciasis, Onchocerciasis and its control. Report of a WHO Expert Committee on Onchocerciasis Control. *World Health Organ Tech Rep Ser* **852**, 1-104 (1995).

8. N. W. Brattig, Pathogenesis and host responses in human onchocerciasis: impact of Onchocerca filariae and Wolbachia endobacteria. *Microbes and infection / Institut Pasteur* **6**, 113-128 (2004).

9. H. Schulz-Key, M. Karam, Quantitative assessment of microfilariae and adults of Onchocerca volvulus in ethanol-fixed biopsies and nodules. *Trans R Soc Trop Med Hyg* **78**, 157-159 (1984).

10. M. G. Basanez *et al*., River blindness: a success story under threat? *PLoS Med* **3**, e371 (2006).

11. H. Schulz-Key, Observations on the reproductive biology of Onchocerca volvulus. *Acta Leiden* **59**, 27-44 (1990).

12. B. O. Duke, Onchocerciasis. *British medical bulletin* **28**, 66-71 (1972).

13. B. O. Duke, The population dynamics of Onchocerca volvulus in the human host. *Trop Med Parasitol* **44**, 61-68 (1993).

14. M. G. Basáñez, J. Ricárdez-Esquinca, Models for the population biology and control of human onchocerciasis. *Trends Parasitol* **17**, 430-438 (2001).

15. M. G. Basáñez *et al*., Density-dependent processes in the transmission of human onchocerciasis: intensity of microfilariae in the skin and their uptake by the simuliid host. *Parasitology* **108 (Pt 1)**, 115-127 (1994).

16. M. G. Basáñez *et al*., Density-dependent processes in the transmission of human onchocerciasis: relationship between the numbers of microfilariae ingested and successful larval development in the simuliid vector. *Parasitology* **110 (Pt 4)**, 409-427 (1995).

17. M. G. Basáñez *et al*., Density-dependent processes in the transmission of human onchocerciasis: relationship between microfilarial intake and mortality of the simuliid vector. *Parasitology* **113 (Pt 4)**, 331-355 (1996).

18. B. O. Duke, Observations and reflections on the immature stages of Onchocerca volvulus in the human host. *Ann Trop Med Parasitol* **85**, 103-110 (1991).

19. G. S. Nelson, Human onchocerciasis: notes on the history, the parasite and the life cycle. *Ann Trop Med Parasitol* **85**, 83-95 (1991).

20. H. P. Duerr, K. Dietz, D. W. Buttner, H. Schulz-Key, A stochastic model for the aggregation of Onchocerca volvulus in nodules. *Parasitology* **123**, 193-201 (2001).

21. F. H. Budden, Comparative study of ocular onchocerciasis in savannah and rain forest. *Trans R Soc Trop Med Hyg* **57**, 64-70 (1963).

22. J. Anderson, H. Fuglsang, P. J. Hamilton, T. F. de Marshall, Studies on onchocerciasis in the United Cameroon Republic. II. Comparison of onchocerciasis in rain-forest and Sudan-savanna. *Trans R Soc Trop Med Hyg* **68**, 209-222 (1974).

23. A. Prost, A. Rougemont, M. S. Omar, [Epidemiological, clinical and biological characters of savanna and forest onchocerciasis in west Africa.--Critical review and new data (author's tranls)]. *Ann Parasitol Hum Comp* **55**, 347-355 (1980).

24. B. O. Duke, D. J. Lewis, P. J. Moore, Onchocerca-Simulium complexes. I. Transmission of forest and Sudan-savanna strains of Onchocerca volvulus, from Cameroon, by Simulium damnosum from various West African bioclimatic zones. *Ann Trop Med Parasitol* **60**, 318-326 (1966).

25. J. Remme, K. Y. Dadzie, A. Rolland, B. Thylefors, Ocular onchocerciasis and intensity of infection in the community. I. West African savanna. *Trop Med Parasitol* **40**, 340-347 (1989).

26. P. A. Zimmerman *et al*., Onchocerca volvulus DNA probe classification correlates with epidemiologic patterns of blindness. *J Infect Dis* **165**, 964-968 (1992).

27. R. A. Cheke, R. Garms, Indices of onchocerciasis transmission by different members of the Simulium damnosum complex conflict with the paradigm of forest and savanna parasite strains. *Acta Trop* **125**, 43-52 (2013).

28. A. Renz, P. Wenk, The distribution of the microfilariae of Onchocerca volvulus in the different body regions in relation to the attacking behaviour of Simulium damnosum s.l. in the Sudan savanna of northern Cameroon. *Trans R Soc Trop Med Hyg* **77**, 748-752 (1983).

29. J. P. Moreau, A. Prost, J. Prod'hon, [An attempt to normalize the methodology of clinico parasitologic surveys of onchocerciasis in West-Africa (author's transl)]. *Med Trop (Mars)* **38**, 43-51 (1978).

30. J. Remme, O. Ba, K. Y. Dadzie, M. Karam, A force-of-infection model for onchocerciasis and its applications in the epidemiological evaluation of the Onchocerciasis Control Programme in the Volta River basin area. *Bull World Health Organ* **64**, 667-681 (1986).

31. R. C. Collins, A. D. Brandling-Bennett, R. B. Holliman, C. C. Campbell, R. F. Darsie, Parasitological diagnosis of onchocerciasis: comparisons of incubation media and incubation times for skin snips. *Am J Trop Med Hyg* **29**, 35-41 (1980).

32. P. Ngoumou, J. F. Walsh, J. M. Mace, A rapid mapping technique for the prevalence and distribution of onchocerciasis: a Cameroon case study. *Ann Trop Med Parasitol* **88**, 463-474 (1994).

33. L. E. Coffeng *et al.*, Onchocerciasis: the pre-control association between prevalence of palpable nodules and skin microfilariae. *PLoS Negl Trop Dis* **7**, e2168 (2013).

34. A. Prost, J. P. Hervouet, B. Thylefors, [Epidemiologic status of onchocerciasis]. *Bull World Health Organ* **57**, 655-662 (1979).

35. P. Stingl, M. Ross, D. W. Gibson, J. Ribas, D. H. Connor, A diagnostic "patch test" for onchocerciasis using topical diethylcarbamazine. *Trans R Soc Trop Med Hyg* **78**, 254-258 (1984).

36. B. A. Boatin *et al.*, Detection of Onchocerca volvulus infection in low prevalence areas: a comparison of three diagnostic methods. *Parasitology* **125**, 545-552 (2002).

37. P. D. Burbelo, H. P. Leahy, M. J. Iadarola, T. B. Nutman, A four-antigen mixture for rapid assessment of Onchocerca volvulus infection. *PLoS Negl Trop Dis* **3**, e438 (2009).

38. L. S. Ayong *et al.*, Development and evaluation of an antigen detection dipstick assay for the diagnosis of human onchocerciasis. *Trop Med Int Health* **10**, 228-233 (2005).

39. G. J. Weil *et al.*, A rapid-format antibody card test for diagnosis of onchocerciasis. *J Infect Dis* **182**, 1796-1799 (2000).

40. E. M. Lipner *et al.*, Field applicability of a rapid-format anti-Ov-16 antibody test for the assessment of onchocerciasis control measures in regions of endemicity. *J Infect Dis* **194**, 216-221 (2006).

41. A. Golden *et al.*, Extended result reading window in lateral flow tests detecting exposure to Onchocerca volvulus: a new technology to improve epidemiological surveillance tools. *PloS one* **8**, e69231 (2013).

42. R. Morales-Hojas *et al.*, Characterisation of nuclear ribosomal DNA sequences from Onchocerca volvulus and Mansonella ozzardi (Nematoda: Filarioidea) and development of a PCR-based method for their detection in skin biopsies. *International journal for parasitology* **31**, 169-177 (2001).

43. S. Pischke, D. W. Buttner, E. Liebau, P. Fischer, An internal control for the detection of Onchocerca volvulus DNA by PCR-ELISA and rapid detection of specific PCR products by DNA Detection Test Strips. *Trop Med Int Health* **7**, 526-531 (2002).

44. D. Globisch *et al.*, Onchocerca volvulus-neurotransmitter tyramine is a biomarker for river blindness. *Proc Natl Acad Sci U S A* **110**, 4218-4223 (2013).

45. M. G. Basanez *et al.*, A research agenda for helminth diseases of humans: modelling for control and elimination. *PLoS Negl Trop Dis* **6**, e1548 (2012).

46. A. Hoerauf, N. Brattig, Resistance and susceptibility in human onchocerciasis--beyond Th1 vs. Th2. *Trends Parasitol* **18**, 25-31 (2002).

47. T. F. Marshall, J. Anderson, H. Fuglsang, The incidence of eye lesions and visual impairment in onchocerciasis in relationship to the intensity of infection. *Trans R Soc Trop Med Hyg* **80**, 426-434 (1986).

48. J. E. McMahon, S. I. Sowa, G. H. Maude, B. R. Kirkwood, Onchocerciasis in Sierra Leone 3: Relationships between eye lesions and microfilarial prevalence and intensity. *Trans R Soc Trop Med Hyg* **82**, 601-605 (1988).

49. R. Brown, R. Shannon, Prevalence, intensity and ocular manifestations of Onchocerca volvulus infection in Dimbelenge, Zaire. *Ann Soc Belg Med Trop* **69**, 137-142 (1989).

50. K. Y. Dadzie, J. Remme, A. Rolland, B. Thylefors, Ocular onchocerciasis and intensity of infection in the community. II. West African rainforest foci of the vector Simulium yahense. *Trop Med Parasitol* **40**, 348-354 (1989).

51. K. Y. Dadzie, J. Remme, R. H. Baker, A. Rolland, B. Thylefors, Ocular onchocerciasis and intensity of infection in the community. III. West African rainforest foci of the vector Simulium sanctipauli. *Trop Med Parasitol* **41**, 376-382 (1990).

52. J. Anderson, H. Fuglsang, A. al-Zubaidy, Onchocerciasis in Yemen with special reference to sowda. *Trans R Soc Trop Med Hyg* **67**, 30-31 (1973).

53. D. W. Buttner, P. Racz, Macro- and microfilariae in nodules from onchocerciasis patients in the Yemen Arab Republic. *Tropenmed Parasitol* **34**, 113-121 (1983).

54. A. Hoerauf *et al*., The variant Arg110Gln of human IL-13 is associated with an immunologically hyper-reactive form of onchocerciasis (sowda). *Microbes and infection / Institut Pasteur* **4**, 37-42 (2002).

55. E. B. Ovuga, D. O. Okello, J. W. Ogwal-Okeng, N. Orwotho, R. O. Opoka, Social and psychological aspects of onchocercal skin disease in Nebbi district, Uganda. *East Afr Med J* **72**, 449-453 (1995).

56. W. R. Brieger, F. O. Oshiname, O. O. Ososanya, Stigma associated with onchocercal skin disease among those affected near the Ofiki and Oyan Rivers in western Nigeria. *Soc Sci Med* **47**, 841-852 (1998).

57. O. U. Manafa, A. N. Isamah, Local knowledge and attitudes about onchocerciasis in Oji-River local government area of Enugu State, Nigeria. *Epidemiology and infection* **129**, 629-633 (2002).

58. I. C. Okoye, C. O. Onwuliri, Epidemiology and psycho-social aspects of onchocercal skin diseases in northeastern Nigeria. *Filaria J* **6**, 15 (2007).

59. A. Saint Andre *et al*., The role of endosymbiotic Wolbachia bacteria in the pathogenesis of river blindness. *Science* **295**, 1892-1895 (2002).

60. I. Gillette-Ferguson *et al*., Wolbachia-induced neutrophil activation in a mouse model of ocular onchocerciasis (river blindness). *Infection and immunity* **72**, 5687-5692 (2004).

61. K. Daehnel *et al*., Filaria/Wolbachia activation of dendritic cells and development of Th1-associated responses is dependent on Toll-like receptor 2 in a mouse model of ocular onchocerciasis (river blindness). *Parasite immunology* **29**, 455-465 (2007).

62. M. J. Taylor, A. Hoerauf, Wolbachia bacteria of filarial nematodes. *Parasitol Today* **15**, 437-442 (1999).

63. B. Thylefors, Ocular onchocerciasis. *Bull World Health Organ* **56**, 63-73 (1978).

64. J. O'Day, C. D. Mackenzie, Ocular onchocerciasis. Diagnosis and current clinical approaches. *Trop Doct* **15**, 87-94 (1985).

65. C. S. Ament, L. H. Young, Ocular manifestations of helminthic infections: onchocersiasis, cysticercosis, toxocariasis, and diffuse unilateral subacute neuroretinitis. *Int Ophthalmol Clin* **46**, 1-10 (2006).

66. M. Boussinesq, S. D. Pion, N. Demanga, J. Kamgno, Relationship between onchocerciasis and epilepsy: a matched case-control study in the Mbam Valley, Republic of Cameroon. *Trans R Soc Trop Med Hyg* **96**, 537-541 (2002).

67. C. Kaiser, S. Pion, M. Boussinesq, Head nodding syndrome and river blindness: a parasitologic perspective. *Epilepsia* **50**, 2325-2326 (2009).

68. S. D. Pion *et al*., Epilepsy in onchocerciasis endemic areas: systematic review and meta-analysis of population-based surveys. *PLoS Negl Trop Dis* **3**, e461 (2009).

69. A. B. Raper, R. G. Ladkin, Endemic dwarfism in Uganda. *East Afr Med J* **27**, 339-359 (1950).

70. W. Kipp, G. Burnham, J. Bamuhiiga, M. Leichsenring, The Nakalanga syndrome in Kabarole District, Western Uganda. *Am J Trop Med Hyg* **54**, 80-83 (1996).

71. A. Prost, J. Vaugelade, [Excess mortality among blind persons in the West African savannah zone]. *Bull World Health Organ* **59**, 773-776 (1981).

72. S. D. Pion, J. Kamgno, N. Demanga, M. Boussinesq, Excess mortality associated with blindness in the onchocerciasis focus of the Mbam Valley, Cameroon. *Ann Trop Med Parasitol* **96**, 181-189 (2002).

73. B. Kirkwood, P. Smith, T. Marshall, A. Prost, Relationships between mortality, visual acuity and microfilarial load in the area of the Onchocerciasis Control Programme. *Trans R Soc Trop Med Hyg* **77**, 862-868 (1983).

74. A. Prost, The burden of blindness in adult males in the savanna villages of West Africa exposed to onchocerciasis. *Trans R Soc Trop Med Hyg* **80**, 525-527 (1986).

75. B. O. Duke, J. Anderson, H. Fuglsang, The Onchocerca volvulus transmission potentials and associated patterns of onchocerciasis at four Cameroon Sudan-savanna villages. *Tropenmed Parasitol* **26**, 143-154 (1975).

76. J. Anderson, H. Fuglsang, T. F. Marshall, Studies on onchocerciasis in the United Cameroon Republic. III. A four year follow-up of 6 rain-forest and 6 Sudan-savanna villages. *Trans R Soc Trop Med Hyg* **70**, 362-373 (1977).

77. A. W. Woodruff, J. Anderson, L. E. Pettitt, M. Tukur, A. W. Woodruff, Some aspects of onchocerciasis in Sudan savanna and rain-forest. *J Trop Med Hyg* **80**, 68-73 (1977).

78. J. J. Picq, J. P. Albert, [Sudan-savanna and rain-forest onchocerciasis in West Africa: an epidemiological problem (author's transl)]. *Rev Epidemiol Sante Publique* **27**, 483-498 (1979).

79. J. Romeo De Leon, B. O. Duke, Experimental studies on the transmission of Guatemalan and West African strains of Onchocerca volvulus by Simulium ochraceum, S. metallicum and S. callidum. *Trans R Soc Trop Med Hyg* **60**, 735-752 (1966).

80. T. B. Higazi *et al*., Wolbachia endosymbiont levels in severe and mild strains of Onchocerca volvulus. *Molecular and biochemical parasitology* **141**, 109-112 (2005).

81. S. E. Meredith, G. Lando, A. A. Gbakima, P. A. Zimmerman, T. R. Unnasch, Onchocerca volvulus: application of the polymerase chain reaction to identification and strain differentiation of the parasite. *Experimental parasitology* **73**, 335-344 (1991).

82. P. Fischer, J. Bamuhiiga, A. H. Kilian, D. W. Buttner, Strain differentiation of Onchocerca volvulus from Uganda using DNA probes. *Parasitology* **112 (Pt 4)**, 401-408 (1996).

83. A. J. Shelley, Vector aspects of the epidemiology of onchocerciasis in Latin America. *Annu Rev Entomol* **33**, 337-366 (1988).

84. World Health Organization (WHO), *Onchocerciasis (river blindness) : report from the fourteenth InterAmerican Conference on Onchocerciasis, Atlanta, Georgia, United States*. (2005).

85. S. West, B. Munoz, A. Sommer, River blindness eliminated in Colombia. *Ophthalmic Epidemiol* **20**, 258-259 (2013).

86. B. Boatin, The Onchocerciasis Control Programme in West Africa (OCP). *Ann Trop Med Parasitol* **102 Suppl 1**, 13-17 (2008).

87. K. Y. Dadzie *et al*., Onchocerciasis control by large-scale ivermectin treatment. *Lancet* **337**, 1358-1359 (1991).

88. J. M. Hougard, E. S. Alley, L. Yameogo, K. Y. Dadzie, B. A. Boatin, Eliminating onchocerciasis after 14 years of vector control: a proved strategy. *J Infect Dis* **184**, 497-503 (2001).

89. L. Yameogo, Special intervention zones. *Ann Trop Med Parasitol* **102 Suppl 1**, 23-24 (2008).

90. World Health Organization (WHO), *Success in Africa : the Onchocerciasis Control Programme in West Africa, 1974-2002*. (2002).

91. A. Seketeli *et al*., The achievements and challenges of the African Programme for Onchocerciasis Control (APOC). *Ann Trop Med Parasitol* **96 Suppl 1**, S15-28 (2002).

92. J. Okeibunor *et al*., A model for evaluating the sustainability of community-directed treatment with ivermectin in the African Program for Onchocerciasis Control. *The International journal of health planning and management* **27**, 257-271 (2012).

93. M. Homeida *et al*., APOC's strategy of community-directed treatment with ivermectin (CDTI) and its potential for providing additional health services to the poorest populations. African Programme for Onchocerciasis Control. *Ann Trop Med Parasitol* **96 Suppl 1**, S93-104 (2002).

94. D. H. Molyneux, J. B. Davies, Onchocerciasis control: Moving towards the millennium. *Parasitol Today* **13**, 418-425 (1997).

95. S. E. Meredith, H. B. Dull, Onchocerciasis: The First Decade of Mectizantrade mark Treatment. *Parasitol Today* **14**, 472-474 (1998).

96. J. B. Davies, Sixty years of onchocerciasis vector control: a chronological summary with comments on eradication, reinvasion, and insecticide resistance. *Annu Rev Entomol* **39**, 23-45 (1994).

97. M. A. Rodríguez-Pérez *et al.*, Rapid suppression of Onchocerca volvulus transmission in two communities of the Southern Chiapas focus, Mexico, achieved by quarterly treatments with Mectizan. *Am J Trop Med Hyg* **79**, 239-244 (2008).

98. J. M. Hougard *et al.*, Criteria for the selection of larvicides by the Onchocerciasis Control Programme in west Africa. *Ann Trop Med Parasitol* **87**, 435-442 (1993).

99. J. M. Hougard, L. Yameogo, A. Seketeli, B. Boatin, K. Y. Dadzie, Twenty-two years of blackfly control in the onchocerciasis control programme in West Africa. *Parasitol Today* **13**, 425-431 (1997).

100. C. Lévêque, in *Ecotoxicology and climate : with special reference to hot and cold climates,* P. Bourdeau, J. A. Haines, W. Klein, C. R. Krishna Murti, Eds. (John Wiley & Sons Ltd, Chichester, 1989), chap. 5.

101. M. A. Aziz, S. Diallo, I. M. Diop, M. Lariviere, M. Porta, Efficacy and tolerance of ivermectin in human onchocerciasis. *Lancet* **2**, 171-173 (1982).

102. A. T. White *et al.*, Controlled trial and dose-finding study of ivermectin for treatment of onchocerciasis. *J Infect Dis* **156**, 463-470 (1987).

103. M. G. Basáñez *et al.*, Effect of single-dose ivermectin on Onchocerca volvulus: a systematic review and meta-analysis. *The Lancet infectious diseases* **8**, 310-322 (2008).

104. C. Bottomley, V. Isham, R. C. Collins, M. G. Basanez, Rates of microfilarial production by Onchocerca volvulus are not cumulatively reduced by multiple ivermectin treatments. *Parasitology* **135**, 1571-1581 (2008).

105. S. Klager, J. A. Whitworth, R. J. Post, D. C. Chavasse, M. D. Downham, How long do the effects of ivermectin on adult Onchocerca volvulus persist? *Trop Med Parasitol* **44**, 305-310 (1993).

106. A. P. Plaisier *et al.*, Irreversible effects of ivermectin on adult parasites in onchocerciasis patients in the Onchocerciasis Control Programme in West Africa. *J Infect Dis* **172**, 204-210 (1995).

107. H. C. Turner *et al.*, The cost of annual versus biannual community-directed treatment of onchocerciasis with ivermectin: Ghana as a case study. *PLoS Negl Trop Dis* **7**, e2452 (2013).

108. H. C. Turner *et al.*, Reaching the London Declaration on Neglected Tropical Diseases goals for onchocerciasis: an economic evaluation of increasing the frequency of ivermectin treatment in Africa. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, (2014).

109. H. C. Turner *et al.*, Uncertainty surrounding projections of the long-term impact of ivermectin treatment on human onchocerciasis. *PLoS Negl Trop Dis* **7**, e2169 (2013).

110. T. S. Churcher *et al.*, Identifying sub-optimal responses to ivermectin in the treatment of River Blindness. *Proc Natl Acad Sci U S A* **106**, 16716-16721 (2009).

111. K. Awadzi *et al.*, Thirty-month follow-up of sub-optimal responders to multiple treatments with ivermectin, in two onchocerciasis-endemic foci in Ghana. *Ann Trop Med Parasitol* **98**, 359-370 (2004).

112. K. Awadzi *et al.*, An investigation of persistent microfilaridermias despite multiple treatments with ivermectin, in two onchocerciasis-endemic foci in Ghana. *Ann Trop Med Parasitol* **98**, 231-249 (2004).

113. M. Y. Osei-Atweneboana, J. K. Eng, D. A. Boakye, J. O. Gyapong, R. K. Prichard, Prevalence and intensity of Onchocerca volvulus infection and efficacy of ivermectin in endemic communities in Ghana: a two-phase epidemiological study. *Lancet* **369**, 2021-2029 (2007).

114. E. Cupp, F. Richards, P. Lammie, M. Eberhard, Efficacy of ivermectin against Onchocerca volvulus in Ghana. *Lancet* **370**, 1123; author reply 1124-1125 (2007).

115. J. H. Remme, U. Amazigo, D. Engels, A. Barryson, L. Yameogo, Efficacy of ivermectin against Onchocerca volvulus in Ghana. *Lancet* **370**, 1123-1124; author reply 1124-1125 (2007).

116. World Health Organization: Expert Committee on Onchocerciasis, EXPERT committee on onchocerciasis; first report. *World Health Organ Tech Rep Ser* **87**, 1-37 (1954).

117. B. Philippon, A. Seketeli, [The beginning of the OPC and the status of onchocerciasis in western Africa before control measures]. *Sante* **8**, 27-31 (1998).

118. J. P. McMahon, A review of the control of Simulium vectors of onchocerciasis. *Bull World Health Organ* **37**, 415-430 (1967).

119. World Health Organization: Expert Committee on Onchocerciasis, "WHO Expert Committee on Onchocerciasis [meeting held in Geneva from 21 to 29 April 1986] : third report," *World Health Organization technical report series* (World Health Organization, Geneva, 1987).

120. World Health Organization: Expert Committee on Onchocerciasis, *WHO Expert Committee on Onchocerciasis [meeting held in Geneva from 29 June to 5 July 1965] : second report*. World Health Organization technical report series (World Health Organization, Geneva, 1966), pp. 96 p.

121. G. De Sole *et al.*, Onchocerciasis distribution and severity in five West African countries. *Bull World Health Organ* **69**, 689-698 (1991).

122. P. Ngoumou, J. F. Walsh, WHO Programme for the Prevention of Blindness., UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases., *A manual for rapid epidemiological mapping of onchocerciasis*. (World Health Organization, Geneva, 1993), pp. 37 p.

123. H. R. Taylor, B. O. Duke, B. Munoz, The selection of communities for treatment of onchocerciasis with ivermectin. *Trop Med Parasitol* **43**, 267-270 (1992).

124. African Programme for Onchocerciasis Control (APOC), "Final communiqué of the 11th session of the Joint Action Forum (JAF) of APOC, Paris, France, 6–9 December 2005," (APOC, Ouagadougou (Burkina Faso), 2005).

125. L. E. Coffeng *et al*., African Programme For Onchocerciasis Control 1995-2015: model-estimated health impact and cost. *PLoS Negl Trop Dis* **7**, e2032 (2013).

126. P. J. Hotez *et al*., Control of neglected tropical diseases. *N Engl J Med* **357**, 1018-1027 (2007).

127. P. J. Hotez, M. E. Bottazzi, C. Franco-Paredes, S. K. Ault, M. R. Periago, The neglected tropical diseases of latin america and the Caribbean: a review of disease burden and distribution and a roadmap for control and elimination. *PLoS Negl Trop Dis* **2**, e300 (2008).

128. P. J. Hotez, A. Fenwick, L. Savioli, D. H. Molyneux, Rescuing the bottom billion through control of neglected tropical diseases. *Lancet* **373**, 1570-1575 (2009).

129. P. J. Hotez, B. Pecoul, "Manifesto" for advancing the control and elimination of neglected tropical diseases. *PLoS Negl Trop Dis* **4**, e718 (2010).

130. P. Diggle, P. J. Ribeiro, SpringerLink (Online service), in *Springer series in statistics,*. (Springer Science+Business Media, LLC,, New York, NY, 2007).

131. D. A. Karagiannis-Voules, R. G. Scholte, L. H. Guimaraes, J. Utzinger, P. Vounatsou, Bayesian geostatistical modeling of leishmaniasis incidence in Brazil. *PLoS Negl Trop Dis* **7**, e2213 (2013).

132. S. Brooker, Spatial epidemiology of human schistosomiasis in Africa: risk models, transmission dynamics and control. *Trans R Soc Trop Med Hyg* **101**, 1-8 (2007).

133. S. Brooker, A. C. Clements, Spatial heterogeneity of parasite co-infection: Determinants and geostatistical prediction at regional scales. *International journal for parasitology* **39**, 591-597 (2009).

134. R. J. Magalhaes, A. C. Clements, A. P. Patil, P. W. Gething, S. Brooker, The applications of model-based geostatistics in helminth epidemiology and control. *Adv Parasitol* **74**, 267-296 (2011).

135. N. A. Wardrop *et al*., Bayesian geostatistical analysis and prediction of Rhodesian human African trypanosomiasis. *PLoS Negl Trop Dis* **4**, e914 (2010).

136. A. P. Patil, P. W. Gething, F. B. Piel, S. I. Hay, Bayesian geostatistics in health cartography: the perspective of malaria. *Trends Parasitol* **27**, 246-253 (2011).

137. H. Slater, E. Michael, Mapping, bayesian geostatistical analysis and spatial prediction of lymphatic filariasis prevalence in Africa. *PloS one* **8**, e71574 (2013).

138. H. G. Zoure *et al*., The geographic distribution of Loa loa in Africa: results of large-scale implementation of the Rapid Assessment Procedure for Loiasis (RAPLOA). *PLoS Negl Trop Dis* **5**, e1210 (2011).

139. A. S. Barro, T. J. Oyana, Predictive and epidemiologic modeling of the spatial risk of human onchocerciasis using biophysical factors: a case study of Ghana and Burundi. *Spatial and spatio-temporal epidemiology* **3**, 273-285 (2012).

140. M. Winnen *et al*., Can ivermectin mass treatments eliminate onchocerciasis in Africa? *Bull World Health Organ* **80**, 384-391 (2002).

141. H. P. Duerr, C. C. Leary, M. Eichner, High infection rates at low transmission potentials in West African onchocerciasis. *International journal for parasitology* **36**, 1367-1372 (2006).

142. M. G. Basáñez, R. C. Collins, C. H. Porter, M. P. Little, D. Brandling-Bennett, Transmission intensity and the patterns of Onchocerca volvulus infection in human communities. *Am J Trop Med Hyg* **67**, 669-679 (2002).

143. J. F. Walsh, J. B. Davies, R. Le Berre, R. Grams, Standardization of criteria for assessing the effect of Simulium control in onchocerciasis control programmes. *Trans R Soc Trop Med Hyg* **72**, 675-676 (1978).

144. African Programme for Onchocerciasis Control, "Conceptual and operational framework of onchocerciasis elimination with ivermectin treatment," (WHO/APOC, 2010).

145. M. J. Taylor *et al*., Onchocerciasis Control: Vision for the Future from a Ghanian perspective. *Parasit Vectors* **2**, 7 (2009).

146. World Health Organization: Expert Committee on Onchocerciasis, Epidemiology of onchocerciasis. Report of a WHO Expert Committee. *World Health Organ Tech Rep Ser*, 1-94 (1976).

147. A. Prost, B. Thylefors, C. Pairault, "Methods of mass epidemiological evaluation of onchocerciasis. Their utilisation in a vector control programme," (World Health Organization, Geneva, 1975).

148. B. Kirkwood, P. Smith, T. Marshall, A. Prost, Variations in the prevalence and intensity of microfilarial infections by age, sex, place and time in the area of the Onchocerciasis Control Programme. *Trans R Soc Trop Med Hyg* **77**, 857-861 (1983).

149. A. Prost, J. Prod'hon, [Parasitological diagnosis of onchocerciasis. A critical review of present methods (author's transl)]. *Med Trop (Mars)* **38**, 519-532 (1978).

150. R. Le Berre, R. Garms, J. B. Davies, J. F. Walsh, B. Philippon, Displacements of Simulium damnosum and strategy of control against onchocerciasis. *Philos Trans R Soc Lond B Biol Sci* **287**, 277-288 (1979).

151. A. Renz, P. Wenk, Studies on the dynamics of transmission of onchocerciasis in a Sudan-savanna area of North Cameroon I. Prevailing Simulium vectors, their biting rates and age-composition at different distances from their breeding sites. *Ann Trop Med Parasitol* **81**, 215-228 (1987).

152. FAO GEONETWORK, Food and Agriculture Organization of the United Nations, Ed. (2014), chap. 1994.

153. A. N. STRAHLER, DYNAMIC BASIS OF GEOMORPHOLOGY. *Geological Society of America Bulletin* **63**, 923-938 (1952).

154. M. G. Basanez, T. S. Churcher, M. E. Grillet, Onchocerca-Simulium interactions and the population and evolutionary biology of Onchocerca volvulus. *Adv Parasitol* **68**, 263-313 (2009).

155. P. H. Adler, R. A. Cheke, R. J. Post, Evolution, epidemiology, and population genetics of black flies (Diptera: Simuliidae). *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **10**, 846-865 (2010).

156. ESRI. (Environmental Systems Research Institute., Redlands, CA, 2012).

157. A. Okabe, A. S. t. Okabe, *Spatial tessellations : concepts and applications of Voronoi diagrams*. (Wiley, Chichester, ed. 2nd ed., Atsuyuki Okabe ... [et al.], 2000).

158. W. S. Cleveland, Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association* **74**, 829-836 (1979).

159. N. Mantel, The detection of disease clustering and a generalized regression approach. *Cancer research* **27**, 209-220 (1967).

160. P. Legendre, M. J. Fortin, Spatial Pattern and Ecological Analysis. *Vegetatio* **80**, 107-138 (1989).

161. S. Holm, A Simple Sequentially Rejective Multiple Test Procedure. *Scand J Stat* **6**, 65-70 (1979).

162. D. Pfeiffer, *Spatial analysis in epidemiology*. (Oxford University Press, Oxford ; New York, 2008), pp. xii, 142 p.

163. T. Jombart, adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403-1405 (2008).

164. K. R. Gabriel, R. R. Sokal, A New Statistical Approach to Geographic Variation Analysis. *Systematic Zoology* **18**, 259-278 (1969).

165. G. T. Toussaint, The Relative Neighborhood Graph of a Finite Planar Set. *Pattern Recogn* **12**, 261-268 (1980).

166. P. A. Moran, Notes on continuous stochastic phenomena. *Biometrika* **37**, 17-23 (1950).

167. L. Anselin, Local Indicators of Spatial Association - Lisa. *Geogr Anal* **27**, 93-115 (1995).

168. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J Roy Stat Soc B Met* **57**, 289-300 (1995).

169. W. N. Venables, B. D. Ripley, W. N. M. a. s. w. S. P. Venables, *Modern applied statistics with S*. (Springer, New York, ed. 4th ed., 2002).

170. G. Wahba, *Spline models for observational data*. (Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1990).

171. S. N. Wood, Thin plate regression splines. *J Roy Stat Soc B* **65**, 95-114 (2003).

172. S. Lee, G. Wolberg, S. Y. Shin, Scattered data interpolation with multilevel B-splines. *Ieee T Vis Comput Gr* **3**, 228-244 (1997).

173. D. Shepard, paper presented at the Proceedings of the 1968 23rd ACM national conference, 1968.

174. N. A. C. Cressie, *Statistics for spatial data*. (Wiley, New York ; Chichester, ed. Rev. ed., 1993).

175. P. J. Diggle, J. A. Tawn, R. A. Moyeed, Model-based geostatistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **47**, 299-350 (1998).

176. G. Matheron, *Traité de géostatistique appliquée*. (Éditions Technip, Paris,, 1962), pp. v.

177. Paulo J. Ribeiro Jr, Peter J. Diggle, in *R-NEWS*. (2001), vol.

1(2):15-18.

178. R Core Team. (R Foundation for Statistical Computing, Vienna, Austria, 2014).

179. Christensen OS, Ribeiro PJ Jr. (R News, 2002), vol. 2: 26-28.

180. N. Cressie, Fitting Variogram Models by Weighted Least-Squares. *J Int Ass Math Geol* **17**, 563-586 (1985).

181. H. Akaike, A new look at the statistical model identification. *Automatic Control, IEEE Transactions on* **19**, 716-723 (1974).

182. African Programme for Onchocerciasis Control (WHO/APOC), "Transforming APOC into a new regional entity for Oncho & LF elimination and support to other PC/NTD," (Ouagadougou, Burkina Faso, 2013).

183. African Programme for Onchocerciasis Control (WHO/APOC), "Programme for the Elimination of Neglected Diseases in Africa (PENDA). Strategic Plan of Action and Indicative Budget 2016-2025.," (Ouagadougou, Burkina Faso, 2013).

184. H. C. Turner, M. Walker, T. S. Churcher, M. G. Basanez, Modelling the impact of ivermectin on River Blindness and its burden of morbidity and mortality in African Savannah: EpiOncho projections. *Parasit Vectors* **7**, 241 (2014).

185. L. Diawara *et al*., Feasibility of onchocerciasis elimination with ivermectin treatment in endemic foci in Africa: first evidence from studies in Mali and Senegal. *PLoS Negl Trop Dis* **3**, e497 (2009).

186. M. O. Traore *et al*., Proof-of-principle of onchocerciasis elimination with ivermectin treatment in endemic foci in Africa: final results of a study in Mali and Senegal. *PLoS Negl Trop Dis* **6**, e1825 (2012).

187. A. H. Tekle *et al*., Impact of long-term treatment of onchocerciasis with ivermectin in Kaduna State, Nigeria: first evidence of the potential for elimination in the operational area of the African Programme for Onchocerciasis Control. *Parasit Vectors* **5**, 28 (2012).

188. World Health Organization., *Accelerating work to overcome the global impact of neglected tropical diseases : a roadmap for implementation : executive summary*. (World Health Organization, Geneva, 2012), pp. 22 p.

189. C. Vajime, D. Quillevere, The distribution of the simulium damnosum complex in West Africa with particular reference to the onchocerciasis control programme area. *Tropenmed Parasitol* **29**, 473-478I (1978).

190. D. A. Boakye, C. Back, G. K. Fiasorgbor, A. P. Sib, Y. Coulibaly, Sibling species distributions of the Simulium damnosum complex in the west African Onchocerciasis Control Programme area during the decade 1984-93, following intensive larviciding since 1974. *Med Vet Entomol* **12**, 345-358 (1998).

191. R. J. Hijmans, S. E. Cameron, J. L. Parra, P. G. Jones, A. Jarvis, Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* **25**, 1965-1978 (2005).

192. J. J. Danielson, paper presented at the Pecora Thirteen, Human Interactions with the Environment – Perspectives from Space, Sioux Falls, South Dakota, August 20–22, 1996. 1996.

193. K. L. Verdin, in *17th Annual Environmental Systems Research Institute (ESRI) User Conference*. (San Diego, California, 1997).

194. C. J. Tucker, J. E. Pinzon, M. E. Brown, Global Land Cover Facility, Ed. (University of Maryland, College Park, Maryland, 2004).

195. J. E. Pinzon, M. E. Brown, C. J. Tucker, in *Hilbert-Huang Transform and its Applications,* N. E. Huang, S. S. P. Shen, Eds. (World Scientific Publishing Co. Pte. Ltd, Singapore, 2005), vol. Interdisciplinary Mathematical Sciences Volume 5, pp. 167-186.

196. C. J. Tucker, J. E. Pinzon, M. E. Brown, D. Slayback, E. W. Pak, et al., An extended AVHRR 8-km NDVI data set compatible with MODIS and SPOT vegetation NDVI data. *Int J Remote Sens* **26**, 4485–4498 (2005).

197. NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS), Ed. (Sioux Falls, South Dakota., 2001).

198. NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS), Ed. (Sioux Falls, South Dakota., 2001).

199. NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS), Ed. (Sioux Falls, South Dakota., 2001).

200. M. A. Friedl *et al*., MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment* **114**, 168-182 (2010).

201. F. WHITE, "Vegetation of Africa - a descriptive memoir to accompany the Unesco/AETFAT/UNSO vegetation map of Africa," (U. N. Educational, Scientific and Cultural Organization (UNESCO), 7 Place de Fontenoy, 75700 Paris, France, 1983).

202. D. B. Gesch, K. S. Larson, in *Pecora Thirteen, Human Interactions with the Environment--Perspectives from Space, 13th*. (Sioux Falls, South Dakota, 1996).

203. U.S. Geological Survey. (Sioux Falls, South Dakota, 1996).

204. GLOBE Task Team *et al*., National Oceanic and Atmospheric Administration, Ed. (National Geophysical Data Center, 325 Broadway, Boulder, Colorado 80305-3328, 1999).

205. Center for International Earth Science Information Network (CIESIN), Columbia University; and Centro Internacional de Agricultura Tropical (CIAT), Socioeconomic Data and Applications Center (SEDAC) Columbia University, Ed. (Palisades, NY, 2005).

206. Center for International Earth Science Information Network (CIESIN) Columbia University; International Food Policy Research Institute (IFPRI); the World Bank; and Centro Internacional de Agricultura Tropical (CIAT), Socioeconomic Data and Applications Center (SEDAC) Columbia University, Ed. (Palisades, NY, 2011).

207. P. C. Austin, J. V. Tu, Bootstrap Methods for Developing Predictive Models. *The American Statistician* **58**, 131-137 (2004).

208. D. J. Spiegelhalter, N. G. Best, B. P. Carlin, A. Van Der Linde, Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**, 583-639 (2002).

209. S. I. Hay *et al*., A world malaria map: Plasmodium falciparum endemicity in 2007. *PLoS Med* **6**, e1000048 (2009).

210. P. W. Gething *et al*., A new world malaria map: Plasmodium falciparum endemicity in 2010. *Malaria journal* **10**, 378 (2011).

211. B. Efron, R. Tibshirani, Improvements on Cross-Validation: The .632+ Bootstrap Method. *Journal of the American Statistical Association* **92**, 548-560 (1997).

212. M. Matsumoto, T. Nishimura, Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Model. Comput. Simul.* **8**, 3-30 (1998).

213. R. J. Hyndman, A. B. Koehler, Another look at measures of forecast accuracy. *International Journal of Forecasting* **22**, 679-688 (2006).

214. R. Moyeed, A. Papritz, An Empirical Comparison of Kriging Methods for Nonlinear Spatial Point Prediction. *Mathematical Geology* **34**, 365-386 (2002).

215. Statisticat LLC. (2013).

216. H. G. Zoure *et al*., The geographic distribution of onchocerciasis in the 20 participating countries of the African Programme for Onchocerciasis Control: (2) pre-control endemicity levels and estimated number infected. *Parasit Vectors* **7**, 326 (2014).

217. M. D. Wilson *et al*., Deforestation and the spatio-temporal distribution of savannah and forest members of the Simulium damnosum complex in southern Ghana and south-western Togo. *Trans R Soc Trop Med Hyg* **96**, 632-639 (2002).

218. World Health Organization/African Programme for Onchocerciasis Control (APOC), "Onchocerciasis control in special intervention zones including Sierra Leone in the OCP area. Joint Programme Committee Report, September 2002," (World Health Organization/African Programme for Onchocerciasis Control,, Ouagadougou (Burkina Faso), 2002).

219. A. Rolland, **G. Balay**, in *Conférence Technique de l'OCCGE*. (OCCGE, sl, 1968), pp. 3 p. multigr.

220. G. De Sole, J. Giese, F. M. Keita, J. Remme, Detailed epidemiological mapping of three onchocerciasis foci in West Africa. *Acta Trop* **48**, 203-213 (1991).

221. R. Garms, J. F. Walsh, J. B. Davies, Studies on the reinvasion of the Onchocerciasis Control Programme in the Volta River Basin by Simulium damnosum s.I. with emphasis on the south-western areas. *Tropenmed Parasitol* **30**, 345-362 (1979).

222. J. F. Walsh, J. B. Davies, R. Garms, Further studies on the reinvasion of the onchocerciasis control programme by Simulium damnosum s.l.: the effects of an extension of control activities into southern Ivory Coast during 1979. *Tropenmed Parasitol* **32**, 269-273 (1981).

223. R. A. Cheke, R. Garms, Reinfestations of the Southeastern Flank of the Onchocerciasis Control Program Area by Windborne Vectors. *Philos T Roy Soc B* **302**, 471-484 (1983).

224. R. H. Baker *et al.*, Progress in controlling the reinvasion of windborne vectors into the western area of the Onchocerciasis Control Programme in West Africa. *Philos Trans R Soc Lond B Biol Sci* **328**, 731-747, discussion 747-750 (1990).

225. B. O. Duke, P. J. Moore, J. Anderson, Studies on factors influencing the transmission of onchocerciasis. VII. A comparison of the Onchocerca volvulus transmission potentials of Simulium damnosum populations in four Cameroon rain-forest villages and the pattern of onchocerciasis associated therewith. *Ann Trop Med Parasitol* **66**, 219-234 (1972).

226. J. B. Davies, P. A. Beech-Garwood, M. C. Thomson, J. E. McMahon, Onchocerciasis transmission levels and Simulium damnosum complex biting activity at riverside and rice field sites in Sierra Leone. *Med Vet Entomol* **2**, 357-369 (1988).

227. M. J. Bockarie, J. B. Davies, M. C. Thomson, H. G. Morgan, The transmission of onchocerciasis at a forest village in Sierra Leone. I. Simulium damnosum s.l. biting densities and infection with Onchocerca volvulus at five representative sites. *Ann Trop Med Parasitol* **84**, 587-597 (1990).

228. J. M. Mace *et al.*, Country-wide rapid epidemiological mapping of onchocerciasis (REMO) in Cameroon. *Ann Trop Med Parasitol* **91**, 379-391 (1997).

229. J. Mendoza Aldana, H. Piechulek, J. Maguire, Forest onchocerciasis in Cameroon: its distribution and implications for selection of communities for control programmes. *Ann Trop Med Parasitol* **91**, 79-86 (1997).

230. S. Vivas-Martinez *et al.*, Onchocerciasis in the Amazonian focus of southern Venezuela: altitude and blackfly species composition as predictors of endemicity to select communities for ivermectin control programmes. *Trans R Soc Trop Med Hyg* **92**, 613-620 (1998).

231. I. Tada *et al.*, Onchocerciasis in San Vicente Pacaya, Guatemala. *Am J Trop Med Hyg* **28**, 67-71 (1979).

232. H. Carabin *et al.*, Prediction of community prevalence of human onchocerciasis in the Amazonian onchocerciasis focus: Bayesian approach. *Bull World Health Organ* **81**, 482-490 (2003).

233. C. Botto *et al.*, in *Encyclopedia of Environmental Health,* J. O. Nriagu, Ed. (Elsevier, Burlington, 2011), pp. 366-379.

234. P. J. Diggle *et al.*, Spatial modelling and the prediction of Loa loa risk: decision making under uncertainty. *Ann Trop Med Parasitol* **101**, 499-509 (2007).

235. F. Chammartin *et al.*, Soil-transmitted helminth infection in South America: a systematic review and geostatistical meta-analysis. *Lancet Infect Dis* **13**, 507-518 (2013).

236. T. Garske, N. M. Ferguson, A. C. Ghani, Estimating air temperature and its influence on malaria transmission across Africa. *PloS one* **8**, e56487 (2013).

237. M. C. Thomson, V. Obsomer, M. Dunne, S. J. Connor, D. H. Molyneux, Satellite mapping of Loa loa prevalence in relation to ivermectin use in west and central Africa. *The Lancet* **356**, 1077-1078 (2000).

238. A. H. Tekle *et al.*, Integrated rapid mapping of onchocerciasis and loiasis in the Democratic Republic of Congo: impact on control strategies. *Acta Trop* **120 Suppl 1**, S81-90 (2011).

239. J. P. Chippaux, M. Boussinesq, J. Gardon, N. Gardon-Wendel, J. C. Ernould, Severe adverse reaction risks during mass treatment with ivermectin in loiasis-endemic areas. *Parasitol Today* **12**, 448-450 (1996).

240. J. Gardon *et al.*, Serious reactions after mass treatment of onchocerciasis with ivermectin in an area endemic for Loa loa infection. *Lancet* **350**, 18-22 (1997).

241. C. J. Murray, A. D. Lopez, D. T. Jamison, The global burden of disease in 1990: summary results, sensitivity analysis and future directions. *Bull World Health Organ* **72**, 495-509 (1994).

242. M. Noma *et al.*, Rapid epidemiological mapping of onchocerciasis (REMO): its application by the African Programme for Onchocerciasis Control (APOC). *Ann Trop Med Parasitol* **96 Suppl 1**, S29-39 (2002).

243. C. J. L. Murray, A. D. Lopez, Harvard School of Public Health., World Health Organization., World Bank., *The global burden of disease : a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*. Global burden of disease and injury series (Published by the Harvard School of Public Health on behalf of the World Health Organization and the World Bank ; Distributed by Harvard University Press, Cambridge, MA, 1996), pp. xxxii, 990 p.

244. C. J. Murray *et al.*, GBD 2010: design, definitions, and metrics. *Lancet* **380**, 2063-2066 (2012).

245. C. J. Murray *et al.*, Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2197-2223 (2012).

246. World Bank. (World Bank., Washington, D.C.), pp. v.

247. T. Arnesen, E. Nord, The value of DALY life: problems with ethics and validity of disability adjusted life years. *Bmj* **319**, 1423-1425 (1999).

248. T. B. Ustun, S. Saxena, J. Rehm, J. Bickenbach, Are disability weights universal? WHO/NIH Joint Project CAR Study Group. *Lancet* **354**, 1306 (1999).

249. J. Jelsma, V. G. Chivaura, K. Mhundwa, W. De Weerdt, P. de Cock, The global burden of disease disability weights. *Lancet* **355**, 2079-2080 (2000).

250. D. D. Reidpath, P. Allotey, Infant mortality rate as an indicator of population health. *J Epidemiol Community Health* **57**, 344-346 (2003).

251. D. D. Reidpath, P. A. Allotey, A. Kouame, R. A. Cummins, Measuring health in a vacuum: examining the disability weight of the DALY. *Health Policy Plan* **18**, 351-356 (2003).

252. J. A. Salomon *et al.*, Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. *Lancet* **380**, 2129-2143 (2012).

253. M. E. Murdoch *et al.*, Onchocerciasis: the clinical and epidemiological burden of skin disease in Africa. *Ann Trop Med Parasitol* **96**, 283-296 (2002).

254. United Nations (UN), UN Department of Economic and Social Affairs Population Division, Ed. (New York, 2013).

255. Amazigo U, Noma M, e. a. Bump J, in *Disease and Mortality in Sub-Saharan Africa. 2nd edition.,* Jamison DT, Feachem RG, Makgoba MW et al, Eds. (World Bank, Washington (DC), 2006), chap. Chapter 15.

256. K. Awadzi, N. O. Opoku, E. T. Addy, B. T. Quartey, The chemotherapy of onchocerciasis. XIX: The clinical and laboratory tolerance of high dose ivermectin. *Trop Med Parasitol* **46**, 131-137 (1995).

257. K. Awadzi, K. Y. Dadzie, S. Klager, H. M. Gilles, The chemotherapy of onchocerciasis. XIII. Studies with ivermectin in onchocerciasis patients in northern Ghana, a region with long lasting vector control. *Trop Med Parasitol* **40**, 361-366 (1989).

258. K. Awadzi *et al.*, The chemotherapy of onchocerciasis X. An assessment of four single dose treatment regimes of MK-933 (ivermectin) in human onchocerciasis. *Ann Trop Med Parasitol* **79**, 63-78 (1985).

259. B. O. Duke, G. Zea-Flores, J. Castro, E. W. Cupp, B. Munoz, Effects of three-month doses of ivermectin on adult Onchocerca volvulus. *Am J Trop Med Hyg* **46**, 189-194 (1992).

260. B. M. Greene *et al.*, Comparison of ivermectin and diethylcarbamazine in the treatment of onchocerciasis. *N Engl J Med* **313**, 133-138 (1985).

261. H. Schulz-Key, M. Karam, Periodic reproduction of Onchocerca volvulus. *Parasitol Today* **2**, 284-286 (1986).

262. P. H. Lamberton *et al.*, Onchocerciasis transmission in Ghana: persistence under different control strategies and the role of the simuliid vectors. *PLoS Negl Trop Dis* **9**, e0003688 (2015).

263. S. D. Pion *et al.*, Dynamics of Onchocerca volvulus microfilarial densities after ivermectin treatment in an ivermectin-naive and a multiply treated population from Cameroon. *PLoS Negl Trop Dis* **7**, e2084 (2013).

264. K. K. Frempong *et al.*, Does Increasing Treatment Frequency Address Suboptimal Responses to Ivermectin for the Control and Elimination of River Blindness? *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, (2016).

265. J. C. Okeibunor *et al.*, Intention to continue with ivermectin treatment for onchocerciasis control after eight years of annual distribution in Cameroon, Nigeria, and Uganda. *International quarterly of community health education* **33**, 159-173 (2012).

266. W. R. Brieger *et al.*, Characteristics of persons who complied with and failed to comply with annual ivermectin treatment. *Trop Med Int Health* **17**, 920-930 (2012).

267. F. Nuwaha, J. Okware, R. Ndyomugyenyi, Predictors of compliance with community-directed ivermectin treatment in Uganda: quantitative results. *Trop Med Int Health* **10**, 659-667 (2005).

268. O. C. Abosi, J. Brookman-Amissah, *Introduction to education in Ghana.* (Sedco, Accra, 1992), pp. x, 292 p.

269. A. Endale, B. Erko, F. Weldegebreal, M. Legesse, Predictors of compliance with community-directed treatment with ivermectin for onchocerciasis control in Kabo area, southwestern Ethiopia. *Parasit Vectors* **8**, 99 (2015).

270. W. R. Brieger *et al.*, Compliance with eight years of annual ivermectin treatment of onchocerciasis in Cameroon and Nigeria. *Parasit Vectors* **4**, 152 (2011).

271. W. A. Oyibo, A. F. Fagbenro-Beyiku, Evaluation of community compliance with annual ivermectin treatment of onchocerciasis in Patigi, Nigeria. *East Afr Med J* **75**, 237-242 (1998).

272. F. Nuwaha, J. Okware, R. Ndyomugyenyi, Predictors for compliance with community directed ivermectin treatment in Bushenyi district of Uganda: qualitative results. *East Afr Med J* **81**, 92-96 (2004).

273. J. T. Talbot *et al.*, Predictors of compliance in mass drug administration for the treatment and prevention of lymphatic filariasis in Leogane, Haiti. *Am J Trop Med Hyg* **78**, 283-288 (2008).

274. M. L. Amarillo, V. Y. Belizario, Jr., J. T. Sadiang-Abay, S. A. Sison, A. M. Dayag, Factors associated with the acceptance of mass drug administration for the elimination of lymphatic filariasis in Agusan del Sur, Philippines. *Parasit Vectors* **1**, 14 (2008).

275. D. W. Njomo, M. Amuyunzu-Nyamongo, J. K. Magambo, S. M. Njenga, The role of personal opinions and experiences in compliance with mass drug administration for lymphatic filariasis elimination in Kenya. *PloS one* **7**, e48395 (2012).

276. B. V. Babu, G. R. Babu, Coverage of, and compliance with, mass drug administration under the programme to eliminate lymphatic filariasis in India: a systematic review. *Trans R Soc Trop Med Hyg* **108**, 538-549 (2014).

277. P. T. Cantey, G. Rao, J. Rout, L. M. Fox, Predictors of compliance with a mass drug administration programme for lymphatic filariasis in Orissa State, India 2008. *Trop Med Int Health* **15**, 224-231 (2010).

278. M. Omedo *et al.*, The effect of a health communication campaign on compliance with mass drug administration for schistosomiasis control in western Kenya--the SCORE project. *Am J Trop Med Hyg* **91**, 982-988 (2014).

279. M. Walker *et al.*, Modelling the elimination of river blindness using long-term epidemiological and programmatic data from Mali and Senegal. *Epidemics* **18**, 4-15 (2017).

280. E. K. Thomsen *et al.*, Efficacy, Safety, and Pharmacokinetics of Coadministered Diethylcarbamazine, Albendazole, and Ivermectin for Treatment of Bancroftian Filariasis. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **62**, 334-341 (2016).

281. M. A. Irvine *et al.*, Effectiveness of a triple-drug regimen for global elimination of lymphatic filariasis: a modelling study. *Lancet Infect Dis* **17**, 451-458 (2017).

282. D. H. Molyneux, A. Hopkins, M. H. Bradley, L. A. Kelly-Hope, Multidimensional complexities of filariasis control in an era of large-scale mass drug administration programmes: a can of worms. *Parasit Vectors* **7**, 363 (2014).

283. Expanded Special Project for Elimination of Neglected Tropical Diseases (ESPEN). vol. 2018.

284. J. Cano *et al*., Identifying co-endemic areas for major filarial infections in sub-Saharan Africa: seeking synergies and preventing severe adverse events during mass drug administration campaigns. *Parasit Vectors* **11**, 70 (2018).

285. R. M. Flueckiger *et al*., Integrating data and resources on neglected tropical diseases for better planning: the NTD mapping tool (NTDmap.org). *PLoS Negl Trop Dis* **9**, e0003400 (2015).

286. R. L. Pullan, M. C. Freeman, P. W. Gething, S. J. Brooker, Geographical inequalities in use of improved drinking water supply and sanitation across Sub-Saharan Africa: mapping and spatial analysis of cross-sectional survey data. *PLoS Med* **11**, e1001626 (2014).

287. A. V. Kulinkina *et al*., Indicators of improved water access in the context of schistosomiasis transmission in rural Eastern Region, Ghana. *Sci Total Environ* **579**, 1745-1755 (2017).

288. I. Ruyssen, G. Rayp, Determinants of Intraregional Migration in Sub-Saharan Africa 1980-2000. *The Journal of Development Studies* **50**, 426-443 (2014).

289. P. Deville *et al*., Dynamic population mapping using mobile phone data. *Proc Natl Acad Sci U S A* **111**, 15888-15893 (2014).

290. A. J. Tatem *et al*., Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning. *Malaria journal* **13**, 52 (2014).

291. A. Wesolowski *et al*., Commentary: containing the ebola outbreak - the potential and challenge of mobile network data. *PLoS Curr* **6**, (2014).

292. L. Bengtsson *et al*., Using mobile phone data to predict the spatial spread of cholera. *Sci Rep* **5**, 8923 (2015).

293. A. Wesolowski *et al*., Quantifying seasonal population fluxes driving rubella transmission dynamics using mobile phone data. *Proc Natl Acad Sci U S A* **112**, 11114-11119 (2015).

294. M. Druet-Cabanac *et al*., Onchocerciasis and epilepsy: a matched case-control study in the Central African Republic. *Am J Epidemiol* **149**, 565-570 (1999).

295. C. Kaiser, G. Asaba, M. Leichsenring, G. Kabagambe, High incidence of epilepsy related to onchocerciasis in West Uganda. *Epilepsy Res* **30**, 247-251 (1998).

296. L. E. Coffeng *et al*., Elimination of African onchocerciasis: modeling the impact of increasing the frequency of ivermectin mass treatment. *PloS one* **9**, e115886 (2014).

297. J. Prod'hon, [Onchocerciasis in African countries not included in the onchocerciasis control program]. *Sante* **8**, 43-48 (1998).

298. B. Boatin *et al*., Patterns of epidemiology and control of onchocerciasis in west Africa. *J Helminthol* **71**, 91-101 (1997).

299. K. J. Beven, M. J. Kirkby, N. Schofield, A. F. Tagg, Testing a Physically-Based Flood Forecasting-Model (Topmodel) for 3 Uk Catchments. *J Hydrol* **69**, 119-143 (1984).

Appendix A. Vectors of Onchocerciasis

Taxon	Principal area	Habitat / epidemiology
African and Arabian Peninsula		
***S. damnosum* complex**		
S. damnosum subcomplex		
S. damnosum s.s. (sensu stricto)	West Africa to Southern Sudan and Uganda	Savanna
S. sirbanum	West Africa to Sudan	Dry savanna
S. rasyani	Yemen	Local vector
S. dieguerense	West Africa	Savanna
***S.sanctipauli* subcomplex**		
S. sanctipauli s.s.	West Africa	Forest
S. soubrense	West Africa	Forest
S. leonense	West Africa, especially Guinea, Sierra Leone	Forest
S. konkourense	West Africa	Guinea Highlands
***S. squamosum* subcomplex**		
S. squamosum s.s.	West Africa	Forest and highlands
S. yahense	West Africa	Upland forest
S. mengense	Cameroon	Probable vector in forest habitat
S. kilibanum	Burundi, Malawi, Uganda, Tanzania, Zaire	Local vector in montane habitat
***S. neavei* group**		
S. neavei s.s.	Uganda, Eastern Zaire	Local vector
S. woodi	United Republic of Tanzania	Local vector
S. ethiopense	South-Western Ethiopia	Local vector
S. albivirgutum	Zaire	Local vector
The Americas		
S. ochraceum s.l.	Mexico, Guatemala	Primary vector in highland foci
S. metallicum s.l.	Mexico, Guatemala, northern Venezuela	2^0 vector in C. & S. America; 1^0 vector in Northern Venezuela
S. callidum	Mexico, Guatemala	2^0 vector
S. exigum s.l.	Colombia, Ecuador, northern Venezuela	1^0 vector in Colombia & Ecuador; 2^0 vector in N. Venezuela
S. guianense s.l.	Brazil/Venezuela (Amazonia)	1^0 vector in highlands
S. incrustatum	Brazil/Venezuela (Amazonia)	2^0 vector in highlands
S. oyapockense s.l.	Brazil/Venezuela (Amazonia)	Vector in lowlands
S. quadrivittatum	Central America, Ecuador	2^0 vector in Ecuador
S. limbatum	Brazil/Venezuela (Amazonia)	Suspected vector in highlands

Table A.1: The *Simuliid* vectors involved in onchocerciasis transmission, their geographical distribution and preferred bioclime (reproduced from (*7*)).

Appendix B. Historical Maps of Onchocerciasis Prevalence in West Africa

This appendix reproduces some historical maps of onchocercal knowledge to emphasise the changing nature of knowledge about onchocercal endemicity across time. The distribution maps published by the WHO Expert Committee on Onchocerciasis show almost no changes in distribution and knowledge of disease distribution between 1976 and 1998. Many maps that were based on input data points displayed *mf* prevalence as pie charts overlaid on the maps. The size and scale of the pie charts gives the misleading impression that onchocercal prevalence is fairly homogeneous across the area. From plotting the pre-control data (Figure 2.3 in Chapter 2), it is in fact clear that the distribution of *mf* across the control programme areas was quite heterogeneous. In addition, it is also clear that the observed data represent only a tiny fraction of the overall population at-risk. An unpublished map, from an internal OCP reports (Figure B.7) shows the detailed population distribution for a small portion of the Upper Volta, based on data from the December 1975 census. It is evident that the at-risk population in these areas is greatly dispersed in many villages, and only a comparatively small number could be surveyed.

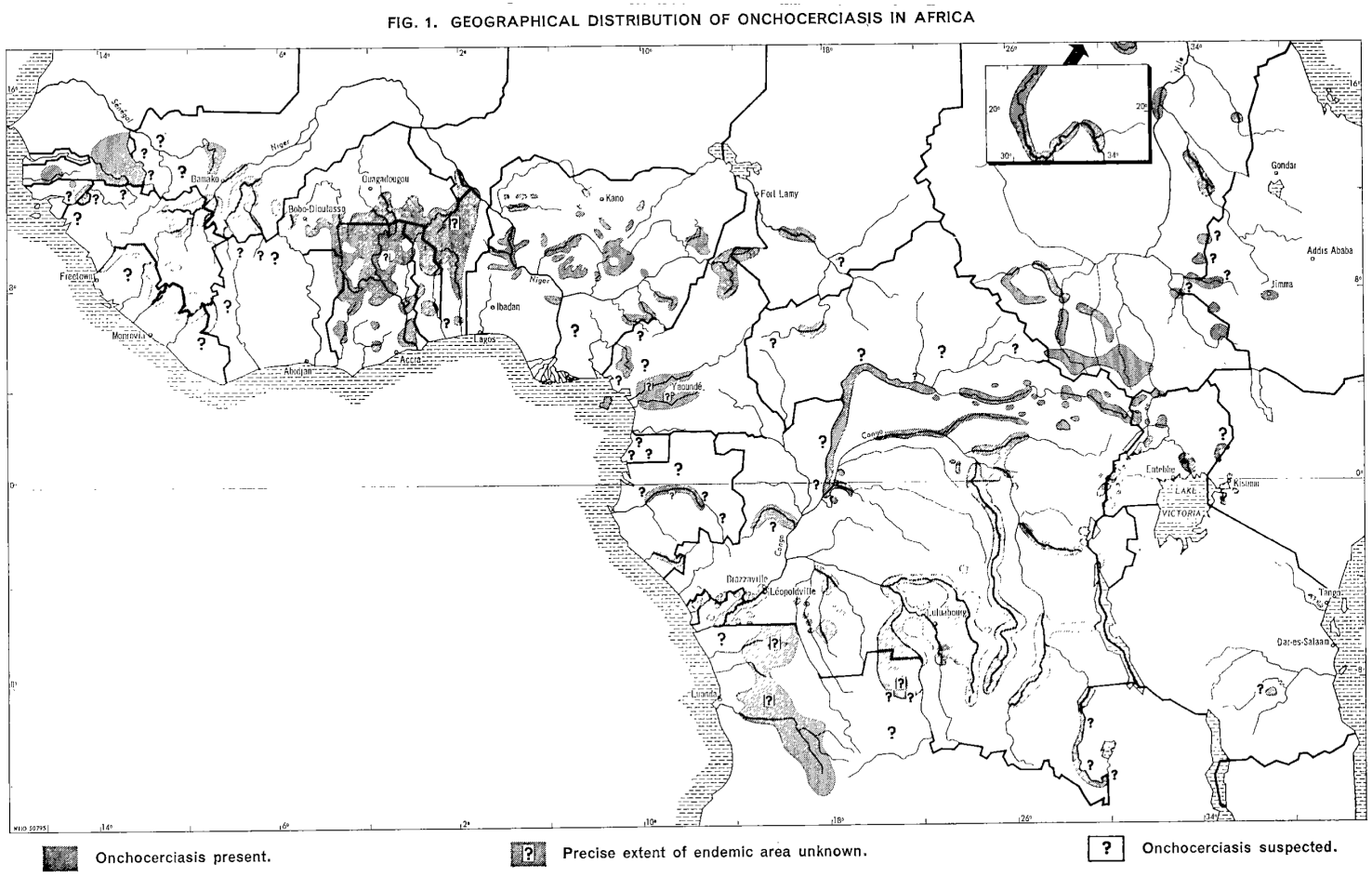
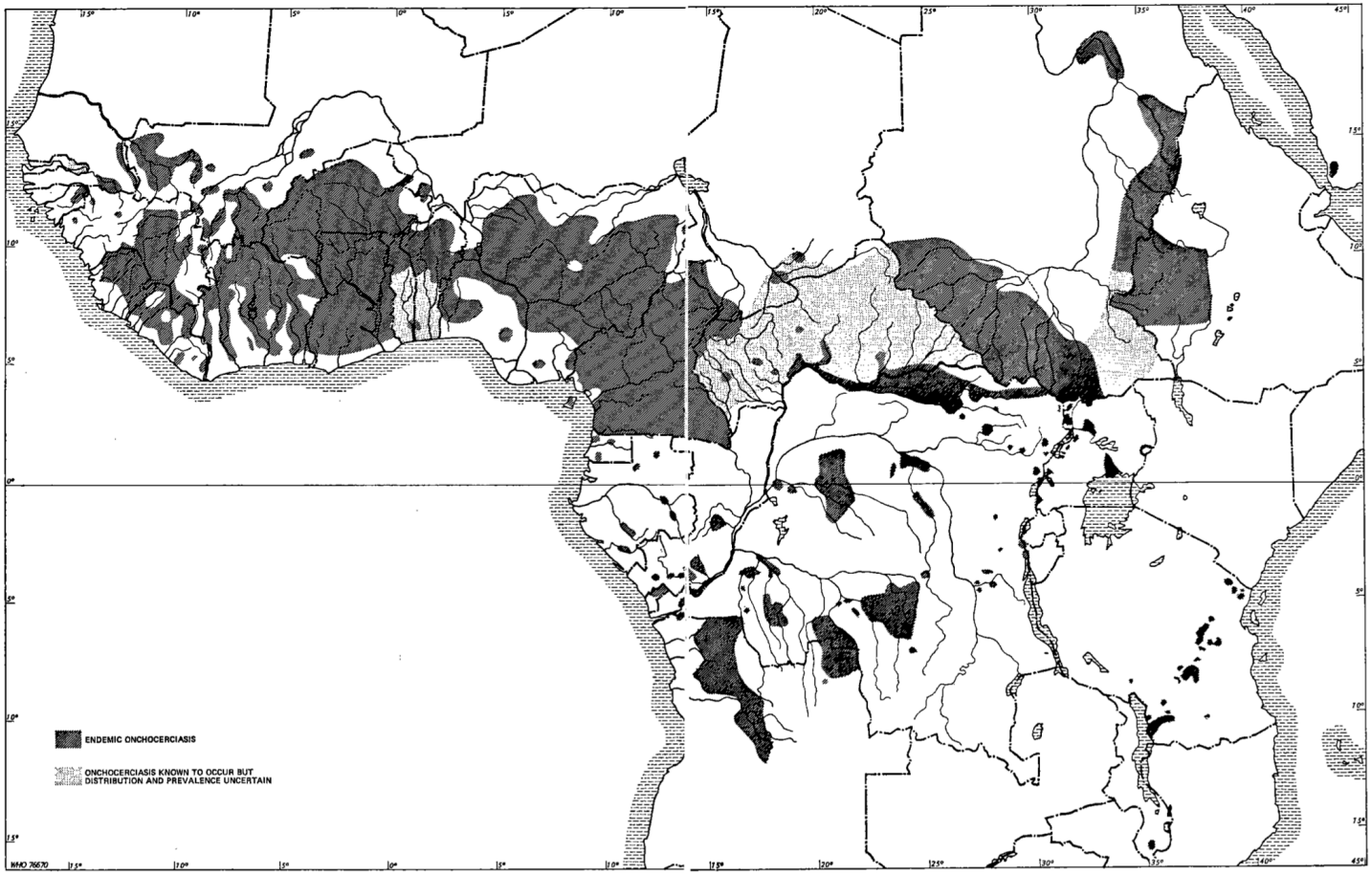


Figure B.1: Onchocerciasis Endemicity in 1966. Map reproduced from Figure 1, WHO, 1966 (*120*).

FIG. 1. GEOGRAPHICAL DISTRIBUTION OF ONCHOCERCIASIS IN THE WORLD: AFRICA AND YEMEN



The map, prepared from information available to WHO at the end of 1975, gives an approximate picture of the main endemic areas as known from publications and reports. White areas do not necessarily indicate absence of disease; they may reflect lack of epidemiological information. Because of the small scale, individual foci within endemic areas are not shown.

8 9

Figure B.2: Estimated Onchocerciasis Endemicity published in 1976. Map reproduced from Figure 1, WHO, 1976 (*146*).

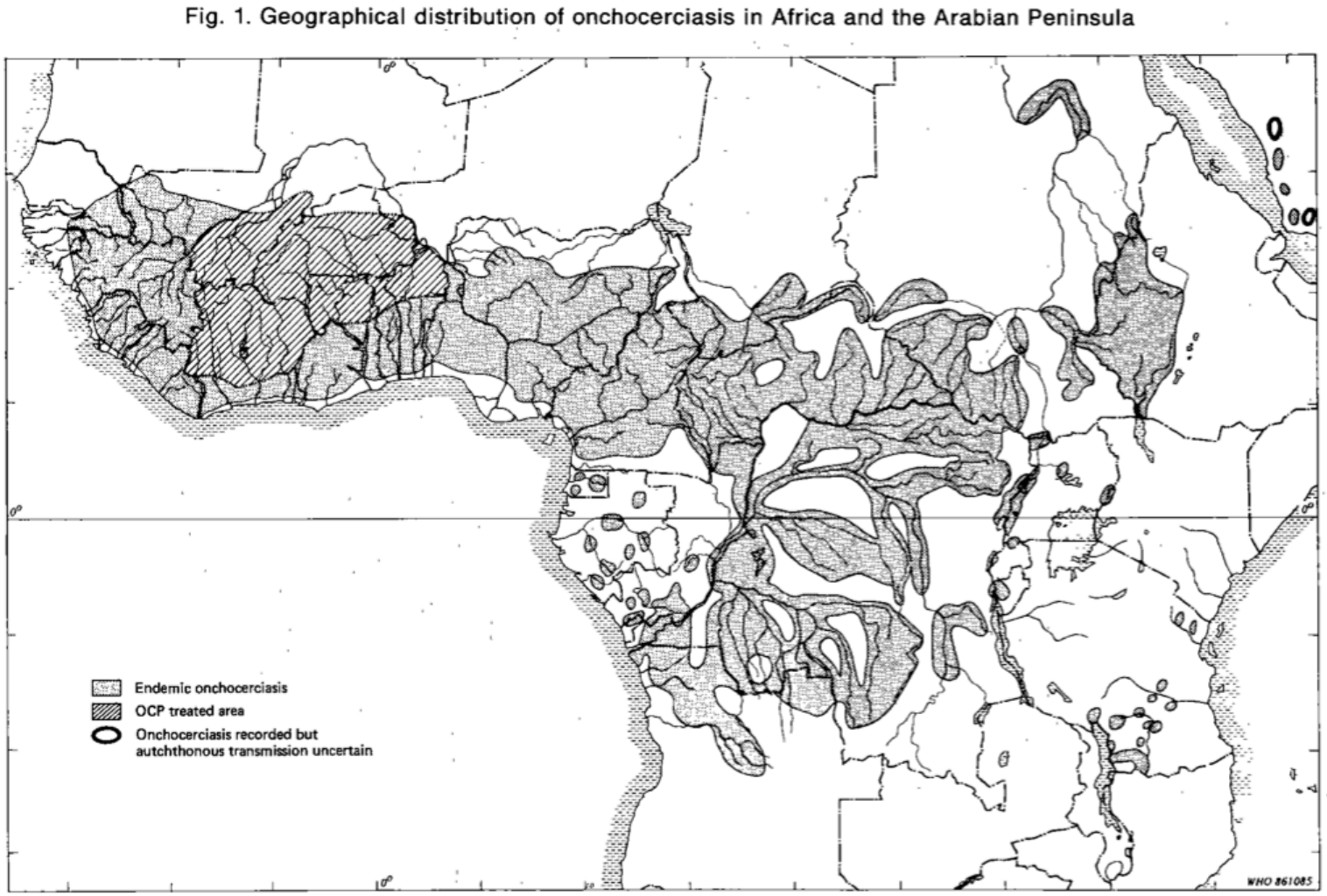


Figure B.3: Estimated Onchocerciasis Endemicity published in 1987. Map reproduced from Figure 1, WHO, 1987 (*119*).

Figure 1
Geographical distribution of onchocerciasis in Africa and the Arabian peninsula

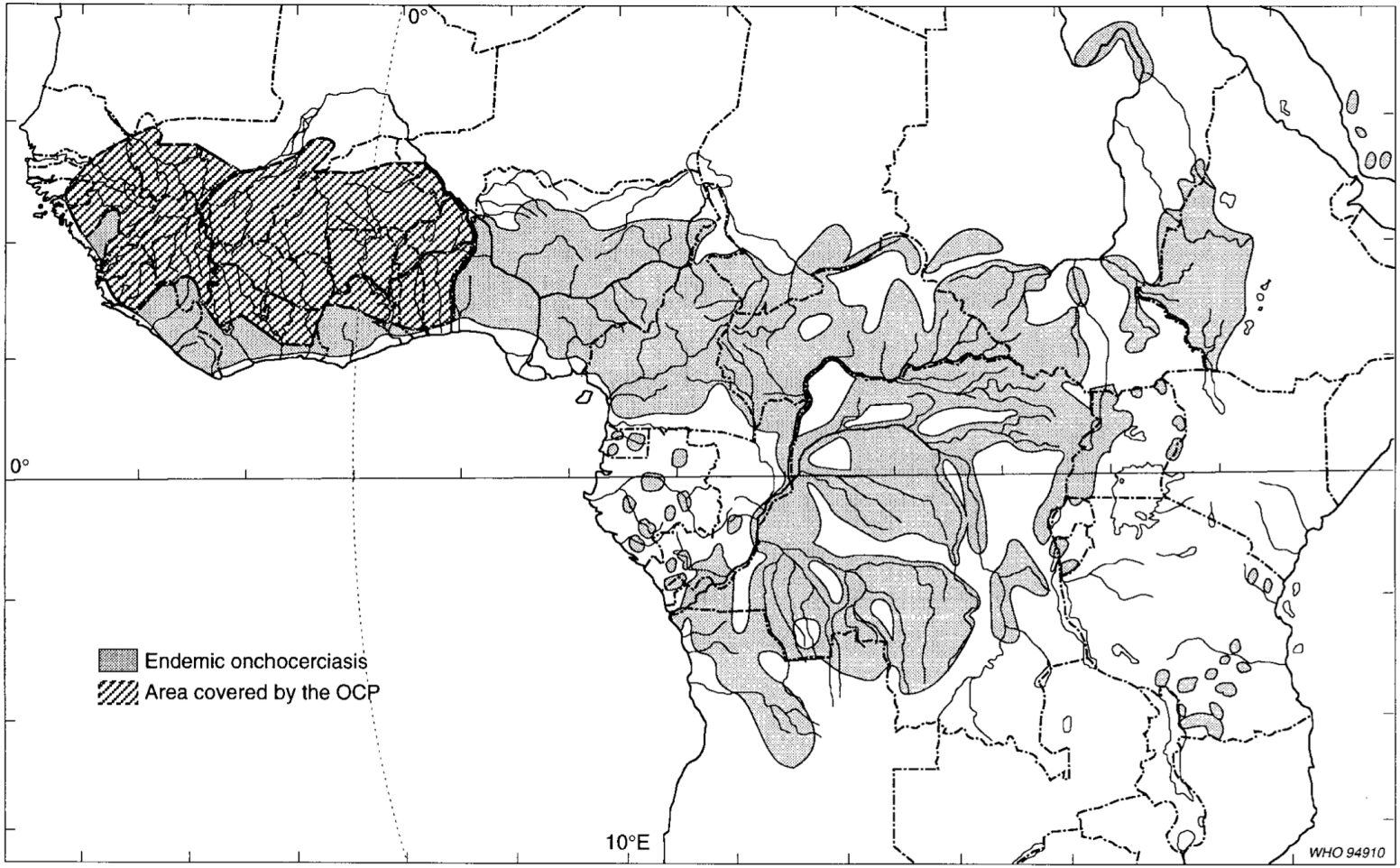


Figure B.4: Estimated Onchocerciasis Distribution published in 1995. Map reproduced from Figure 1, published in WHO (1995) (*7*).

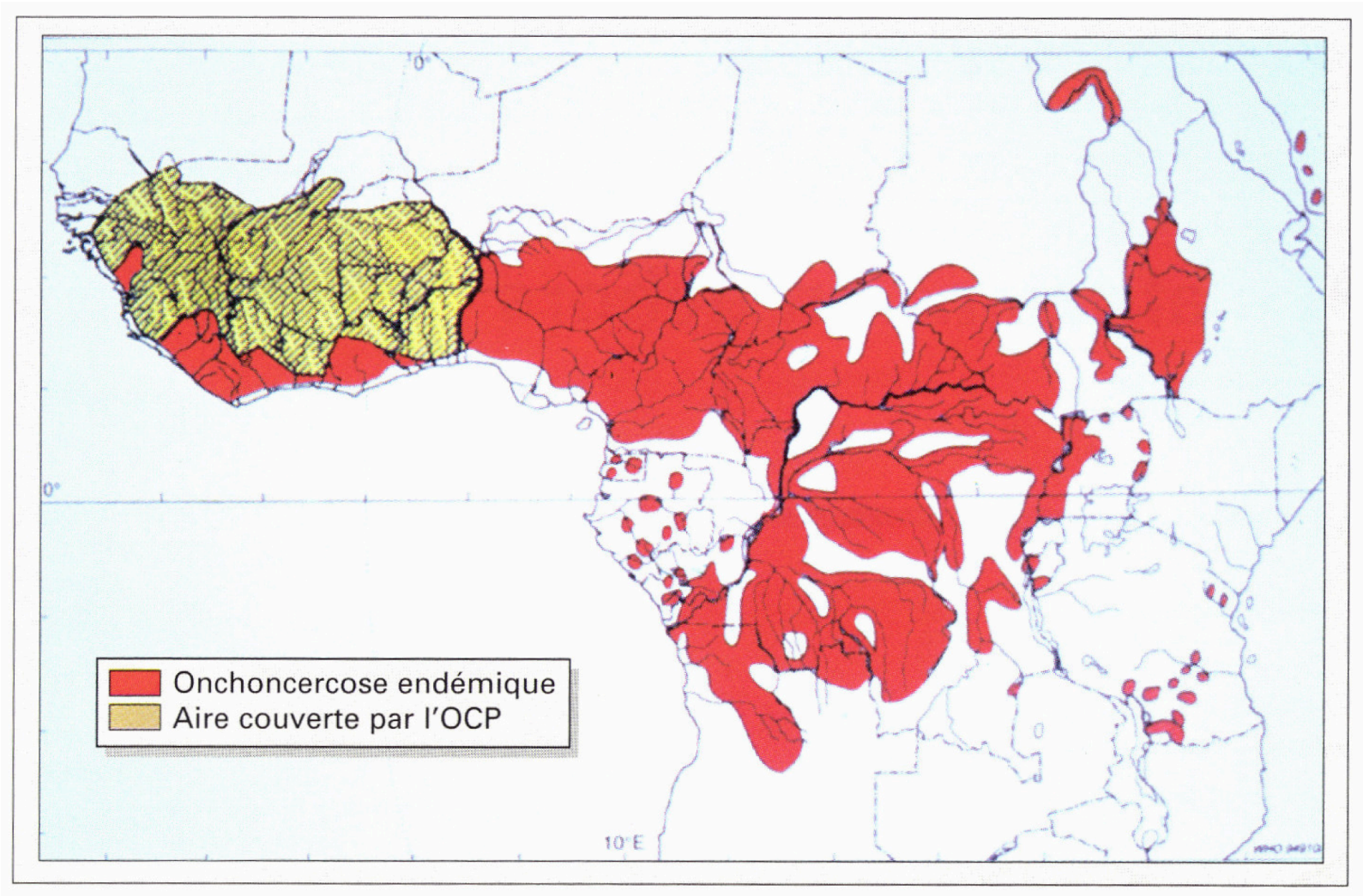


Figure B.5: Estimated Onchocerciasis Distribution in West Africa published in 1998. Reproduced from Figure 1 by Prod'hon, J. (1998) (*297*).

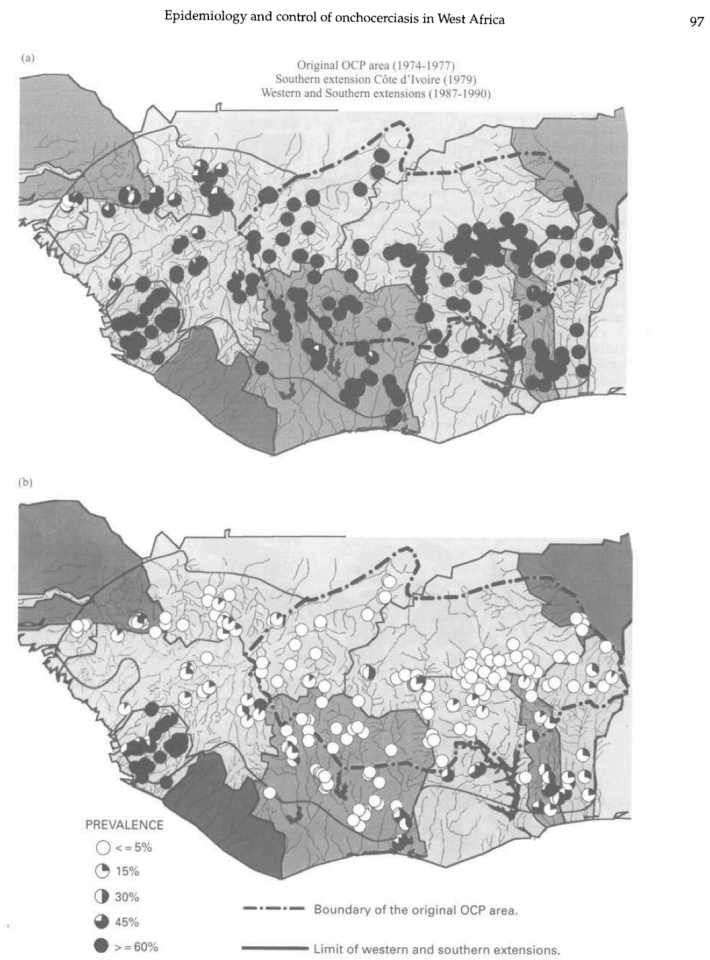


Fig. 4. Prevalence of skin microfilariae pre-control (a) and from 1993 to 1996(b).

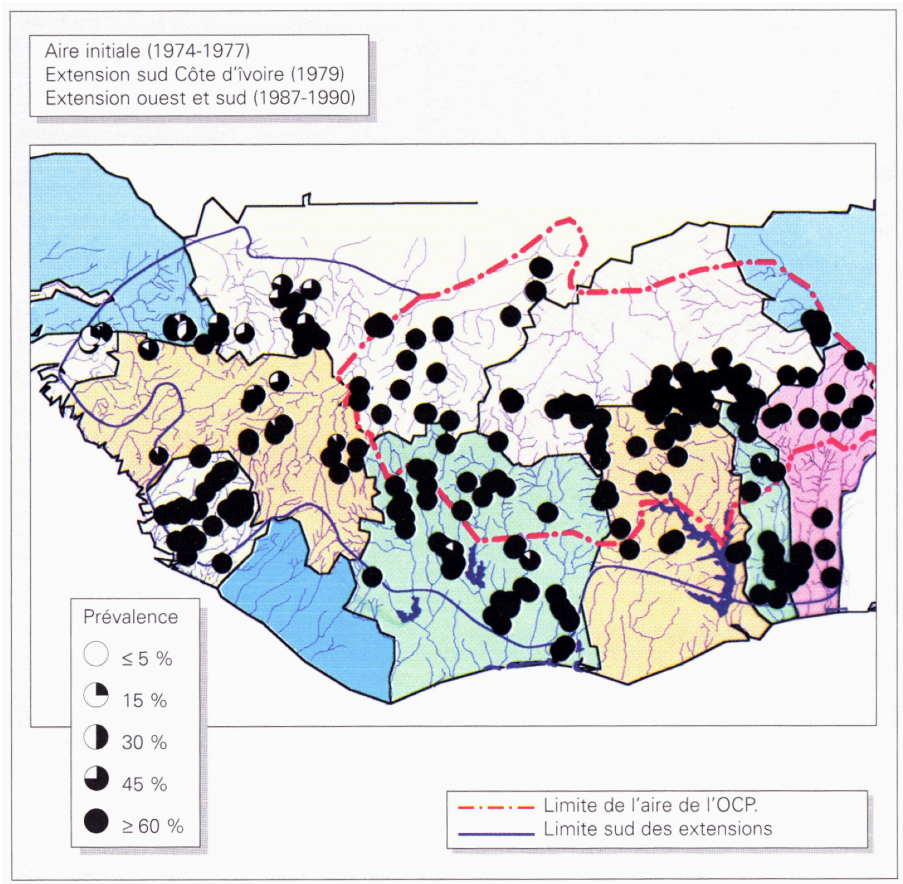


Figure B.6: Pre-control onchocerciasis mf prevalence pie charts. Figure A) was reproduced from Figure 4 in Boatin, B., et al. (1997) (*298*). Figure B) was reproduced from Philippon and Seketeli (1998) (*117*).

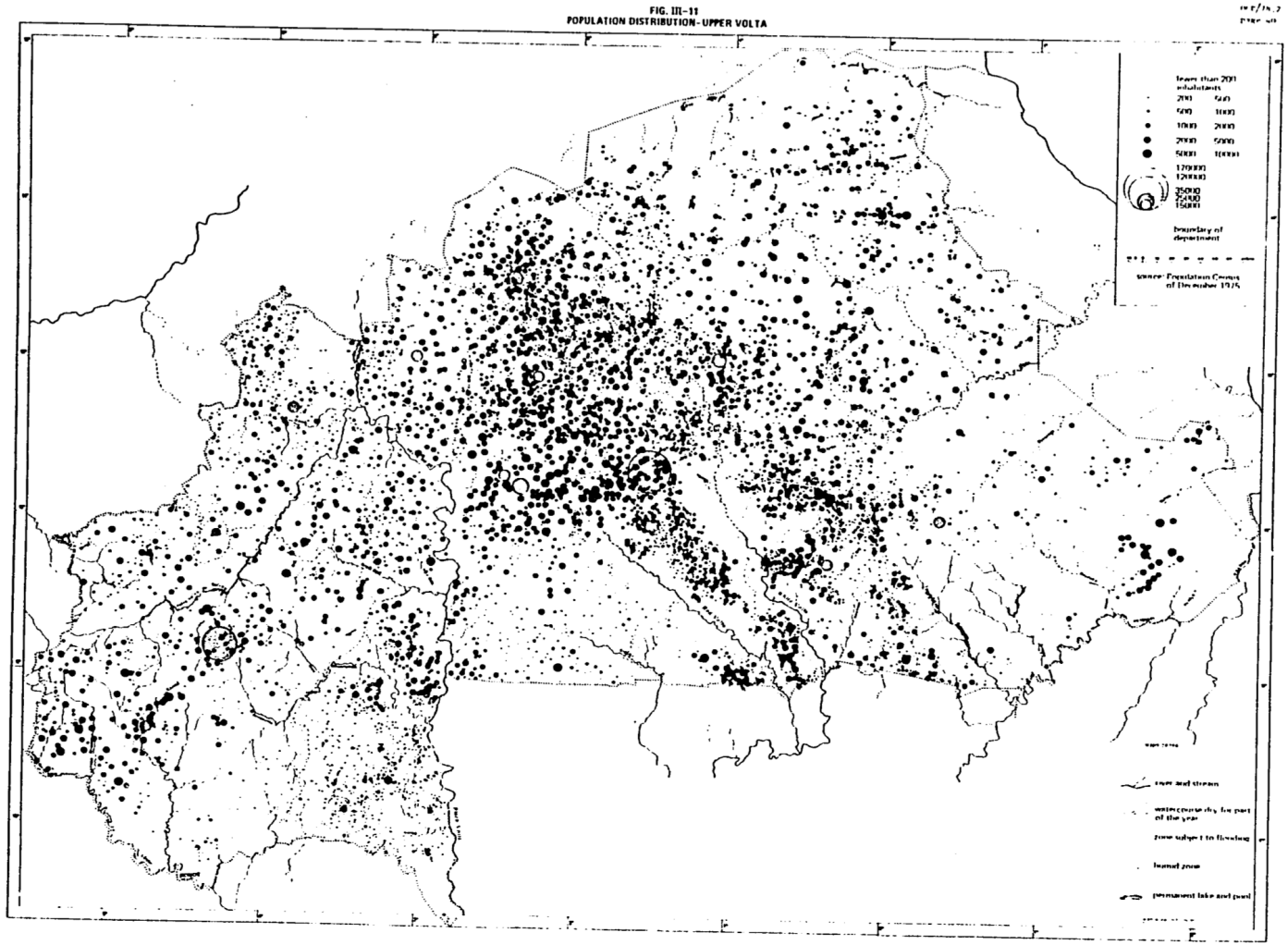


Figure B.7: Detailed population distribution in the Upper Volta, from December 1975 census data. Figure reproduced from Figure 3 from an unpublished OCP report, compiled by the WHO (1978), entitled "Onchocerciasis Control Programme in the Volta River Basin Area. Evaluation Report. Part I." WHO Document number OCP/78.2.

Appendix C. Processing of environmental covariates

An ensemble of potential explanatory environmental covariates was assembled from a variety of data sources, at a variety of spatial resolutions and differing map projections. Selection of environmental covariates was guided by those deemed to have a known influence on the transmission dynamics of *Onchocerca volvulus*. In addition, values for covariates had to be available at each of the prediction locations in the study area, with data ideally available at a spatial resolution of 5-Km or better. The full list of environmental covariates that were assessed is in Table 2 in the Main Text. In this section, I describe the data and the spatial re-projection and sampling that was carried out to generate values of environmental covariates at input data locations and continuous maps across prediction locations. For the biological relevance of including any of these covariates in a predictive model for onchocerciasis prevalence, see the *Discussion* in Chapter 3.

C.1. Sampling covariate layers at data locations

Sampling of environmental covariate values at input data locations was done in the native projection of each covariate layer to minimise error introduced by distortion from re-projecting a regular grid of raster data, usually in a geographic coordinate system, to a projected coordinate system. The coordinates of the data locations were instead re-projected (there is no distortion in the re-projection of a single point in space from one geographic coordinate system to another) to match the projection of each covariate layer in turn. Values were then sampled from each covariate layer at the data locations, before the input data locations and all covariate layers were finally transformed to the common projection of the prediction grid.

The projection of the prediction grid chosen was the Lambert Azimuthal Equal Area projection, which is especially suitable for spatial data across a continental-scale and regional scale. One of the benefits is the accurate representation of distances and area at the expense of angular representation (Figure S9). In addition the datum of the projection, typically in metres or kilometres, is easier to interpret in terms of model outputs than decimal degrees. The specifications of the standard prediction grid were:

```
rows : 303, columns: 460, ncells: 139380

resolution (Km): 5, 5  (x, y)

extent: -4000, -1700, -15, 1500  (xmin, xmax, ymin, ymax)

proj4string: +proj=laea +lon_0=20 +lat_0=5 +ellps=sphere
+to_meter=1e3
```

The total number of cells in the prediction grid included cells that overlay water-bodies, the sea and urban-extents (as defined using the GRUMPv1.0 data below). We did not predict in these cells so the total number of prediction points is 112,295.

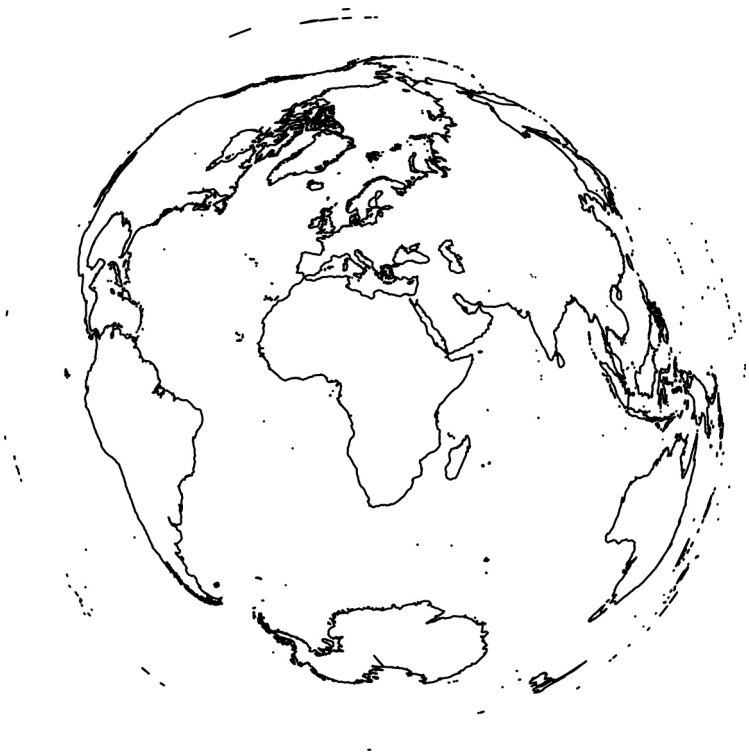


Figure C.1: Lambert Azimuthal Equal Area (LAEA) Projection of the World. The LAEA projection is particularly suited to preserving area, at the expense of angular representation. Accurate representation of distances. This projection is centred on Africa.

C.1.1. WorldClim Data

The WorldClim (*191*) data products are a set of global-coverage raster maps of climatic conditions with a spatial resolution of 1Km. They consist of temporally aggregated data (1950-2000) for average monthly minimum temperature, maximum temperature, mean temperature and precipitation. There are also a set of 19 bioclimatic variables derived from these temperature and precipitation data, which are described at http://www.worldclim.org/bioclim. The temperature data (°C) are supplied as integer data and must be converted to floating point by multiplying by a scaling factor, z, with $z = 0.1$. The units of the precipitation layers are mm. The derived bioclimatic layers (termed Bioclim variables) are more biologically meaningful variables. Examples include precipitation seasonality (the intra-year variability of rainfall, with higher values indicating that rainfall is more seasonal), the mean temperature of the wettest quarter (on a pixel-by-pixel basis, this is the mean temperature for the 12 contiguous weeks that had the highest average rainfall in the year – for neighbouring pixels the 12

contiguous weeks *could* be at different time-points during the year) and mean diurnal range (mean of monthly difference between the maximum temperature and minimum temperature).

C.1.2. Hydro1K Database

Hydrologic covariates were extracted from the Hydro1K database (*192*). These layers consist of several raster maps, derived from a digital elevation model and a vector shapefile of streamlines based on the elevation model. We extracted elevation, stream lines, compound topographic index (a measure of groundwater flow, often called the wetness index), flow direction (direction of flow from each cell in the raster to its steepest down-slope neighbour) and flow accumulation (the upstream catchment area for a given pixel). We used the stream lines data to calculate a raster of Euclidean distance to the nearest stream segment, and derived a raster of gradient of the nearest stream segment as a proxy for flow speed of the nearest river segment. Blackfly larvae survive in rivers with discharge volumes in the range of 0.1-150 cubic metres per second, and stream gradient affects flow rate. The compound topographic index (CTI) or topographic wetness index layer, quantifies how topographic variation affects hydrological processes. Ground water flow often follows surface topography [9], which can also affect river discharge volumes and flow rates, and hence the location of suitable blackfly breeding sites. Following Beven and Kirkby (*299*), the CTI for any given location λ_i is calculated as $\lambda_i = \ln\left(\frac{a_i}{\tan b_i}\right)$ where a_i is the local upstream area which flows into location i, and b_i is the surface slope angle between location i and its neighbouring locations (i.e. the 8 surrounding cells).

C.1.3. MODIS Satellite Data Products

The MOD13Q1 MODIS (*197-200*) data-product contains layers for vegetation indices. We extracted Normalised Difference Vegetation Index & Enhanced Vegetation Index from this dataset. The data are provided at a high resolution of 250m, in 16-day means (i.e. one raster consisting of the mean

values for the previous 16 days). We downloaded all available data (at the time from 2000.02.18 – 2013.07.12), consisting of 309 sets of raster layers. The data are supplied at integer values (to reduce storage and data transmission requirements) and needed to be multiplied by a scaling factor, $z = 0.0001$. We sorted each raster into ecological quarters based on year day of the dataset, with ecological quarters defined as in the Main Text. We then calculated the long-term mean per-pixel across all raster layers in each ecological quarter, resulting in 8 high resolution layers; one for each quarter of NDVI and EVI respectively. After sampling values at input data locations at high resolution, these layers were aggregated to 5km spatial resolution using a bilinear interpolation and re-projected to the same extent and resolution as the prediction grid.

The MOD11A2 data product consists of Land Surface Emissivity data. We used the daytime and nighttime landsurface temperature (LST) layers which were provided at 1Km spatial resolution, as 8-day averages. Temporal aggregation, data sampling, spatial aggregation and re-projection was carried out as above with the exception that the scaling factor $z = 0.02$ and the data were converted from Kelvin to Celsius by subtracting 273.15.

The MCD12Q1 data product contains Yearly Land Cover Type at a spatial resolution of 500m. The data are supplied as a yearly classification with 5 schemes available. We used the International Geosphere-Biosphere Programme classification system [14]. We downloaded data for each of the available years (2001-2012) and aggregated by using the per-pixel mode of land type, suitable for categorical data. This method of aggregation was used in favour of using the earliest available single-layer because there was a period of 26 years between the desired time point (1975) and the earliest available data, and using the mode for each pixel would smooth over any anomalous classification in the earliest raster that was not consistent with long-term use of that pixel. Aggregation from 500-m to 5-km resolution was performed taking the modal value of pixels belonging to the larger pixels.

C.1.4. Global Inventory Modelling and Mapping Studies NDVI

The NDVI layer from the GIMMS database (*194*) is derived from much earlier satellite data from the Advanced Very High Resolution Radiometer (AVHRR) satellites of the National Oceanic and Atmospheric Administration (NOAA). Continuous sampling at an 8-Km resolution is available from 1982 – 2006, however at present the data are listed as 'not available' on the Global Land Cover Facility Website[1]. The data consist of 616 raster layers covering the African continent in twice-monthly averages. The raster data were sampled at input data locations in the native resolution before being resampled. This dataset was downloaded and used because the temporal coverage is much closer to the start date of control of the OCP. However, the low spatial resolution hinders its utility when looking at spatial relationships on a finer scale and it was not found to be as good a univariate predictor of microfilarial prevalence as the high-resolution MODIS data (models using GIMMS NDVI has a higher DIC value in model testing than MODIS NDVI or EVI and were highly correlated with these layers), therefore it was discarded from further analysis.

C.1.5. HydroSheds

The Rivers of Africa dataset is derived from HydroSheds (*152*), a mapping product of the World Wildlife Fund (WWF). The data are based on NASA's Shuttle Radar Topographic Mission (SRTM) 15-second Digital Elevation Model (DEM) and consist of a drainage direction layer and a stream network layer. The stream network consists of a network of spatial lines with attribute information for each line segment. In addition to the spatial location of the stream segments, attributes which were thought to be relevant to this mapping study were the classification of stream segments as perennial or intermittent streams (intermittent streams do not flow throughout the year), and the Strahler order

[1]*The candidate holds a copy of the processed GIMMS NDVI database for the African continent, covering the period from, July 1981- November 2006. I would be pleased to freely provide the dataset upon request on a case-by-case basis, provided I am given the explicit permission of the Global Land Cover Facility, who originally hosted the data.*

(hydrological rank) of river segments. Strahler's stream order system [19], proposes that a river or stream segment be given a rank according to the number of upstream tributaries it possesses. Therefore a 1st order stream segment has no upstream tributaries and is known as a headstream. As rank increases so does the size of a river (as more tributaries feed into it), which can give an indication of the magnitude or flow rate of the river in that location compared to the upstream area.

C.1.6. Whites Vegetation Map of Africa

As well as remotely sensed land cover data we also evaluated the use of a map of major vegetation types based on physiognomy, not climate data (*201*). The data in this layer were compiled from existing maps and verified though fieldwork and local expert knowledge. The morphological types and floristic composition of endemic species were used to define vegetation classifications. The data are supplied as a single raster, which was sampled at native resolution (1-Km) before being aggregated and re-projected to the prediction grid using the same methods as the MODIS MCD12Q1 Land Cover dataset.

C.1.7. GTOPO30 and GLOBE DEMs

There are multiple altitude datasets available. We assessed the use of two DEMs, the Global 30 Arc-Second Elevation (GTOPO30) model (*202*) and the Global Land One-km Base Elevation Project (GLOBE) model (*204*). Both the GLOBE and GTOPO30 provide global coverage of topographic relief data at 1-km resolution. After sampling the value of altitude at input data locations the layers were aggregated and re-projected to the attributes of the prediction grid using a bilinear interpolation. As might be expected from the measurement of a relatively invariant covariate, these datasets were highly congruent. At the input data locations, the mean absolute difference in altitude between the two datasets was 5.81 metres with $r^2 = 0.99$. The linear relationship between altitude and prevalence showed almost complete randomness in both datasets, therefore GLOBE was selected because a

greater percentage of land-pixels (57.5% vs. 50.0% in GTOPO30 DEM) were estimated from the Digital Terrain Elevation Data (DTED) with a native resolution of 3-arc seconds (~90m at the equator).

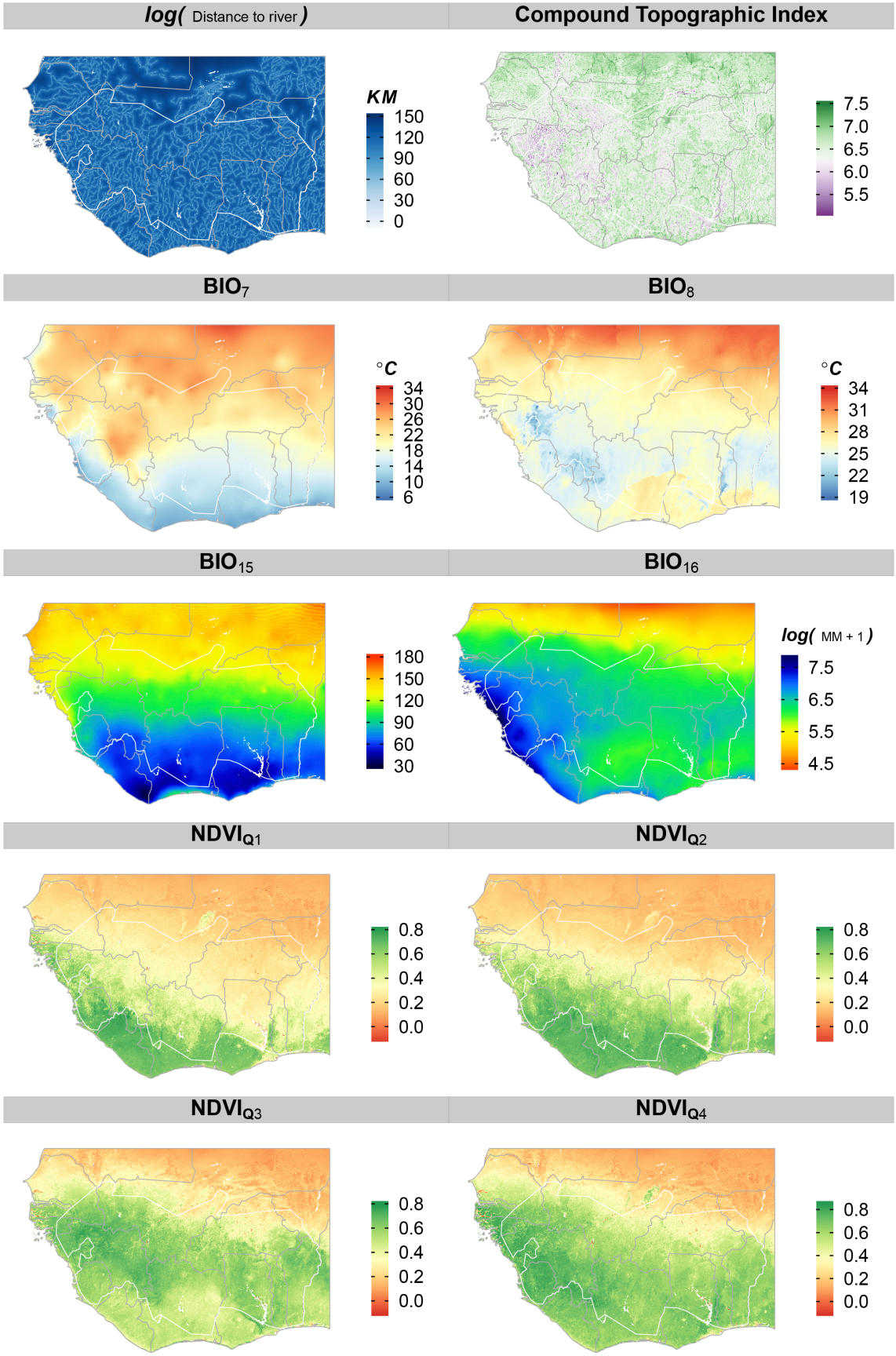


Figure C.2 Maps of covariates included in model from Chapter 3.

Appendix D. Socioeconomic Survey Instrument

A full copy of the socioeconomic survey questionnaire administered to each household in Chapter 5.

INTERNATIONAL DEVELOPMENT RESEARCH CENTER (IDRC)
&
INSTITUTE OF STATISTICAL, SOCIAL & ECONOMIC RESEARCH (ISSER),
UNIVERSITY OF GHANA, LEGON

SOCIO-ECONOMIC SURVEY

ECHO-HEALTH APPROACH TO THE CONTROL OF ONCHOCERCIASIS IN THE VOLTA BASIN OF GHANA

HOUSEHOLD INSTRUMENT, (2010)

REGION:	|_____|	E.A. NO.:	|_____|
DISTRICT:	|_____|	HH NO.:	|_____|
INTERVIEWER ID:	|_____|	SUPERVISOR ID:	|_____|

HOUSEHOLD INFORMATION **Date:** | | | | | | | | |

1. Locality ---------- **2. EAs No**: | | | |

3. Head of Household: ---------- **4. Household No**: | | |

5. Location of household: ----------
---

6. Interviewer ID: | | | **7. Supervisor ID:** | | |

8. Main Languages Spoken at Home:

Asante	11	Guan	41
Fanti	12	Buli	51
Akuapem	13	Mamprusi	52
Sefwi	14	Frafra/Gruni	53
Brong	15	Kassen`	54
Nzema	16	Dagbani	55
Ga	21	Wali/Dagari	56
Dangme	22	Sissala	57
Ewe	31	Other (Speify)	96

9. Type of Housing
1. Separate House (Bungallow)
2. Semi-Detached House
3. Flat (Apartment)
4. Rooms (Compound)
5. Room(s)
6. Room(s) Several Huts/Buildings (Same Compound)
7. Room(s) Several Huts/Buildings (Different Compound)
8. Tents/Improvised Home
9. Other (Specify)

10. How Many Rooms are in this Dwelling (Housing Unit)? | | |

11. Does Another Household Share This Dwelling?
1. Yes
2. No

1

Part A: Consent Form

Title of the Study: Echo-health approach to the control of onchocerciasis in Ghana.
Research Partners: Institute of Statistical, Social and Economic Research (ISSER) at the University of Ghana

My name is _____. I am from the Institute of Statistical, Social and Economic Research at the University of Ghana. I am collecting data for a research study to see how the lives of individuals and households in Ghana are affected by the process of economic change. You have been chosen randomly as one of the participants for this study in Ghana. We expect that the results from this survey will help the Ghanaian government as well as other governments in low-income countries and international organizations to develop policies and procedures that improve the lives of people.

Because many parts of your lives are likely to have been affected by economic change, if you agree to participate in this study we will ask you many detailed questions about such things as who lives in your household, your education and health, your children and others in the household, employment and sources of income, the types and amounts of goods you buy, your sense of the problems in the village as well as the role of various organizations in remedying those problems. The health questions include history of medical care you have received, and how you and your household have been handling issues related to the control of onchocerciasis.

If you choose to take part in the survey, you are free to refuse to answer any of the questions that may make you uncomfortable and you also have the option to end the interview at anytime. You can stop me at anytime to clarify the questions or ask me to repeat something if you don't understand. The question sessions will last for about 1-2 hours at a place of your comfort and convenience. Because we are interested in long term changes in your household and your community we will try to re-contact you next year to conduct follow-up surveys. We realized that your time is valuable and that it may not be easy for you to devote this much time to our survey. We would like you to know that your participation is greatly appreciated and is extremely valuable to the success of our efforts to develop better policies and procedures to improve the lives of Ghanaians and people in other low-income countries.

Your participation in this study will not affect you in anyway as the study is solely for research purposes. The answers you provide will be stored securely in the offices of ISSER and confidentiality will be maintained at all times. Your name and any other identifying information will be accessible only to the researchers and will never appear in any sort of report that might be published.

Your decision to participate or not participate will not affect you. You can contact the research organization that has hired us to conduct this survey at the address below with any problems or questions you may have:

Dr. Isaac Osei-Akoto at the Institute of Statistical, Social and Economic Research University of Ghana, Legon: 021501182

If you **consent** to participate in this study please sign or make a thumb imprint. Your participation in this study is completely voluntary. We are very grateful if you participate. Do you agree to participate in the study?

Yes_____

No _____

Signature (thumbprint if respondent is not a literate) _____ Date: _____

Signature of the Enumerator: _____ Date: _____

1. Are you happy with your life? **1. Yes** **2. No**

Part B1: Future Contact Information

Enumerator: please remind the household that it will be interviewed again in the next 1 year, and thus having a way to contact them in the future is critical to the success of the survey. We are going to ask them for two people who can be contacted in the future should the family move from the village.

1: If you left this place, who would be the most likely people to know where you are?

	Contact 1	**Contact 2**
1. Name of contact people		
2. Relationship to you		
3. Where do these people currently live? 1. Name of community 2. Name of district 3. Name of region	1._____ 2._____ 3. _____	1._____ 2._____ 3. _____
4. Phone numbers of these people		
5. How best would we be able to contact these people? **(Pls. *feel free to give as many options as necessary*)** **e.g. Contact address, landmarks, street numbers, popular names etc**		

Part B2: Household Roster - (To be asked about all household members)

Member ID	*1. Official Names*	*2. Most Popular Name*	*3. Sex (M/F)*	*4. Age (If age less <1, record 0)*	*5. Relationship to HH Head (see codes to the right of the page)*	*6. Marital Status (see codes: 12 years and older*	*7. Current Cell Phone Number*	
1								***CODES FOR Q.5*** *1 – Household Head* *2 – Spouse* *3 – Child* *4 - Grandchild* *5 – Parent/ Parent-in-law* *6 – Son/Daughter-in-law* *7 – Other relative* *8 – Adopted/Foster/ Stepchild* *9 – House help* *10 – Non-relative*
2								
3								
4								
5								
6								
7								
8								
9								***CODES FOR Q.6*** *1 – Never Married* *2 - Consensual union* *3 – Betrothed* *4 – Married* *5 - Separated* *6 – Divorced* *7 - Widowed*
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								

Part C: Background Information - ***(To be asked about all household members)***

Member ID	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1. What is [Name's] date of birth? 1. dd mm yy yy 2.DK	DD MM YY	DD MM YY	DD MM YY	DD MM YY	DD MM YY	DD MM YY	DD MM YY	DD MM YY	DD MM YY	DD MM YY
2. How old is (Name)? 1. yy. mm if <5 ; otherwise, just years.	YY MM	YY MM	YY MM	YY MM	YY MM	YY MM	YY MM	YY MM	YY MM	YY MM
3. Is (Name) <12 years 1. Yes >> 7 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No
4. What is [Name's] marital status? 1. Married - MA 2. Consensual union - CU 3.Separated (>> 7) - SE 4. Divorced (>> 7) - DI 5. Widowed (>> 7) - WI 6. Never married (>> 7) - NM 7.Betrothed - BE	1. MA 2. CU 3. SE 4. DI 5. WI 6. NM 7. BE	1. MA 2. CU 3. SE 4. DI 5. WI 6. NM 7. BE	1. MA 2. CU 3. SE 4. DI 5. WI 6. NM 7. BE	1. MA 2. CU 3. SE 4. DI 5. WI 6. NM 7. BE	1. MA 2. CU 3. SE 4. DI 5. WI 6. NM 7. BE	1. MA 2. CU 3. SE 4. DI 5. WI 6. NM 7. BE	1. MA 2. CU 3. SE 4. DI 5. WI 6. NM 7. BE	1. MA 2. CU 3. SE 4. DI 5. WI 6. NM 7. BE	1. MA 2. CU 3. SE 4. DI 5. WI 6. NM 7. BE	1. MA 2. CU 3. SE 4. DI 5. WI 6. NM 7. BE
5. Does spouse live in the household? 1. yes 2. no (>>7)	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No
6. Spouse's ID (If more than one spouse, get ID of the first spouse) 1. ID	1. |_|_| ID	1. |_|_| ID	1. |_|_| ID	1. |_|_| ID	1. |_|_| ID	1. |_|_| ID	1. |_|_| ID	1. |_|_| ID	1. |_|_| ID	1. |_|_| ID
7. What is [Name's] religious denomination? 1. Catholic 2. Anglican 3. Presbyterian 4. Methodist 5. Pentecostals 6. Spiritualist 7. Other Christian 8. Muslim 9. Traditional 10. No Religion 11.Other non-Christian (specify)	1 2 3 4 5 6 7 8 9 10 11 _____	1 2 3 4 5 6 7 8 9 10 11 _____	1 2 3 4 5 6 7 8 9 10 11 _____	1 2 3 4 5 6 7 8 9 10 11 _____	1 2 3 4 5 6 7 8 9 10 11 _____	1 2 3 4 5 6 7 8 9 10 11 _____	1 2 3 4 5 6 7 8 9 10 11 _____	1 2 3 4 5 6 7 8 9 10 11 _____	1 2 3 4 5 6 7 8 9 10 11 _____	1 2 3 4 5 6 7 8 9 10 11 _____

Member ID	1	2	3	4	5	6	7	8	9	10
8. In what region/country was (Name) born? 1. Western 2. Central 3. Gt. Accra 4. Volta 5. Eastern 6. Ashanti 7. Brong Ahafo 8. Northern 9. Upper East 10. Upper West 96. Other ECOWAS 97. Africa other than ECOWAS 98. Outside of Africa	1 2 3 4 5 6 7 8 9 10 96 97 98	1 2 3 4 5 6 7 8 9 10 96 97 98	1 2 3 4 5 6 7 8 9 10 96 97 98	1 2 3 4 5 6 7 8 9 10 96 97 98	1 2 3 4 5 6 7 8 9 10 96 97 98	1 2 3 4 5 6 7 8 9 10 96 97 98	1 2 3 4 5 6 7 8 9 10 96 97 98	1 2 3 4 5 6 7 8 9 10 96 97 98	1 2 3 4 5 6 7 8 9 10 96 97 98	1 2 3 4 5 6 7 8 9 10 96 97 98
9. What is [Name's] Nationality? 1. Ghanaian (birth) 2. Ghanaian (naturalized) 3. Burkinabe 4. Malian 5. Nigerian 6. Ivorian 7. Togolese 8. Liberian 9. Other ECOWAS 10. Other African 11. Other (Specify)	1 2 3 4 5 6 7 8 9 10 11	1 2 3 4 5 6 7 8 9 10 11	1 2 3 4 5 6 7 8 9 10 11	1 2 3 4 5 6 7 8 9 10 11	1 2 3 4 5 6 7 8 9 10 11	1 2 3 4 5 6 7 8 9 10 11	1 2 3 4 5 6 7 8 9 10 11	1 2 3 4 5 6 7 8 9 10 11	1 2 3 4 5 6 7 8 9 10 11	1 2 3 4 5 6 7 8 9 10 11
10. To what ethnic group does (Name) belong? Ethnicity (write in) See code book and use specifics	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____
11. Has (Name) ever held a political office? 1. None – NO>>13 2. Member of parliament - MP 3. District chief executive – DC 4. Assembly man/woman -AP 5. Unit Committee Member - UC 6. National Political Office (Minister, Deputy Minister ,Political party officials etc.) - NP 7.Other (specify)	1.NO 2.MP 3.DC 4.AP 5.UC 6.NP 7._____	1.NO 2.MP 3.DC 4.AP 5.UC 6.NP 7._____	1.NO 2.MP 3.DC 4.AP 5.UC 6.NP 7._____	1.NO 2.MP 3.DC 4.AP 5.UC 6.NP 7._____	1.NO 2.MP 3.DC 4.AP 5.UC 6.NP 7._____	1.NO 2.MP 3.DC 4.AP 5.UC 6.NP 7._____	1.NO 2.MP 3.DC 4.AP 5.UC 6.NP 7._____	1.NO 2.MP 3.DC 4.AP 5.UC 6.NP 7._____	1.NO 2.MP 3.DC 4.AP 5.UC 6.NP 7._____	1.NO 2.MP 3.DC 4.AP 5.UC 6.NP 7._____

6

Member ID	1	2	3	4	5	6	7	8	9	10
12. Does (Name) still hold this office? **1. Yes** **2. No**	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No
13. Has (Name) ever held a traditional office? 1. None - NO >>15 2. Paramount Chief - PC 3. Sub-chief - SC 4. Queen Mother - QM 5. Head of Family - *Abusupanyin* - HF 6.Chief's Elders/ Linguist - CE 7. Other Palace Officials -PO 8. Community Spirtual Head/ Fetish – CS 9. Other – OT (specify)	1.NO 2.PC 3.SC 4.QM 5.HF 6.CE 7.PO 8.CS 9. OT	1.NO 2.PC 3.SC 4.QM 5.HF 6.CE 7.PO 8.CS 9. OT	1.NO 2.PC 3.SC 4.QM 5.HF 6.CE 7.PO 8.CS 9. OT	1.NO 2.PC 3.SC 4.QM 5.HF 6.CE 7.PO 8.CS 9. OT	1.NO 2.PC 3.SC 4.QM 5.HF 6.CE 7.PO 8.CS 9. OT	1.NO 2.PC 3.SC 4.QM 5.HF 6.CE 7.PO 8.CS 9. OT	1.NO 2.PC 3.SC 4.QM 5.HF 6.CE 7.PO 8.CS 9. OT	1.NO 2.PC 3.SC 4.QM 5.HF 6.CE 7.PO 8.CS 9. OT	1.NO 2.PC 3.SC 4.QM 5.HF 6.CE 7.PO 8.CS 9. OT	1.NO 2.PC 3.SC 4.QM 5.HF 6.CE 7.PO 8.CS 9. OT
14. Does (Name) still hold this office? 1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No
15. For how many months has (Name) been away from this household in last 12 months? 1. Number of months, If less than 6 months (>>Next Person)										
16. If absent, is/was (Name) a member of another household? Including a single person household 1. Yes 2. No	1. yes 2. no	1. yes 2. no	1. yes 2. no	1. yes 2. no	1. yes 2. no	1. yes 2. no	1. yes 2. no	1. yes 2. no	1. yes 2. no	1. yes 2. no
17. For what reason did (Name) move away? 1. Job transfer 2. Seeking employment 3.Spouse's employment 4. Marriage 5. Other family reason 6. Education 7. Political/religious 8. Effects of Oncho disease 9. Effects of other diseases 10. Other (specify)	1. 2 3. 4. 5. 6. 7. 8. 9. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. _____

7

Member ID	1	2	3	4	5	6	7	8	9	10
18. If in Ghana, where did (Name) move? 1. Kumasi - KU 2. Sunyani - SU 3. Cape Coast - CC 4. Koforidua - KO 5. Accra - AC 6. Tamale - TA 7. Bolgatanga - BO 8. Wa - WA 9. Sekondi-Takordi - SK 10. Ho – HO 11. Same village – SV 12. Other village - OV (specify) 13. Other urban – OU (specify)	1.KU 2.SU 3.CC 4.KO 5.AC 6.TA 7.BO 8.WA 9.SK 10.HO 11. SV 12.OV _____ 13. OU _____	1.KU 2.SU 3.CC 4.KO 5.AC 6.TA 7.BO 8.WA 9.SK 10.HO 11. SV 12.OV _____ 13. OU _____	1.KU 2.SU 3.CC 4.KO 5.AC 6.TA 7.BO 8.WA 9.SK 10.HO 11. SV 12.OV _____ 13. OU _____	1.KU 2.SU 3.CC 4.KO 5.AC 6.TA 7.BO 8.WA 9.SK 10.HO 11. SV 12.OV _____ 13. OU _____	1.KU 2.SU 3.CC 4.KO 5.AC 6.TA 7.BO 8.WA 9.SK 10.HO 11. SV 12.OV _____ 13. OU _____	1.KU 2.SU 3.CC 4.KO 5.AC 6.TA 7.BO 8.WA 9.SK 10.HO 11. SV 12.OV _____ 13. OU _____	1.KU 2.SU 3.CC 4.KO 5.AC 6.TA 7.BO 8.WA 9.SK 10.HO 11. SV 12.OV _____ 13. OU _____	1.KU 2.SU 3.CC 4.KO 5.AC 6.TA 7.BO 8.WA 9.SK 10.HO 11. SV 12.OV _____ 13. OU _____	1.KU 2.SU 3.CC 4.KO 5.AC 6.TA 7.BO 8.WA 9.SK 10.HO 11. SV 12.OV _____ 13. OU _____	1.KU 2.SU 3.CC 4.KO 5.AC 6.TA 7.BO 8.WA 9.SK 10.HO 11. SV 12.OV _____ 13. OU _____
19. If (Name) moved somewhere in Ghana, to which district? (use code book)										

8

PART D: EMPLOYMENT
EMPLOYMENT IN LAST 12 MONTHS
COMPLETE FOR ALL HOUSEHOLD MEMBERS AGE 7 OR OLDER

0. Point of verification: In the last 12 months, has anyone in this household done work for pay, profit or for family consumption

1. Yes 2. No >> PART E

Member ID	1	2	3	4	5	6	7	8	9	10
1. Did (name) do any type of work for pay in cash or in kind profit or for family consumption in the last 12 months? **1 Yes >>3** **2 No**										
2. Why is (name) not working/did not do any work? **1. Attending School >> Part E** **2. Too old to work >> 13** **3. Disabled >> 13** **4. Disabled due to Oncho>>13** **5. Lost previous job >> 12** **6. Cannot find a job >> 12** **7. Work is seasonal >> 12** **8. Other (Specify) >> 12**										
3. Have you been looking for work and been ready for work in the last 12 months? **1 Yes** **2 No**										
4. In which economic sector do/did you work in (main job)? **1. Agriculture** **2. Fishing** **3. Mining and quarrying** **4. Manufacturing** **5. Construction** **6. Transport/storage/comm.** **7. Finance/insurance/services** **8. Utilities** **9. Wholesale/retail trade** **10. Community/social services** **11. Other (Specify)**										

Member ID	1	2	3	4	5	6	7	8	9	10
5. What are the problems you face with regard to your work? **1. Finance** **2. Difficult to access land** **3. Erratic weather conditions** **4. Uncertain demand for output** **5. Poor health** **6. Inadequate fish catch** **7. High cost of inputs e.g. fuel** **8. Difficulty in marketing products** **9. Unavailability of inputs** **10. Low price of products** **11. Other (Specify)**										
6. What is your employment status in you main job? **1. Self employed with employees** **2. Self employed without employees** **3. Unpaid family worker** **4. Casual worker** **5. Regular employee** **6. Domestic employee** **7. Apprentice** **8. Other (Specify)**										
7. For whom do you work in the main job? **1. Public** **2. Private formal** **3. Private informal** **4. Semi-public/parastatal** **5. NGO/International organization** **6. Other (Specify)**										
8. How much has (name) earned during the last 12 months from this economic activity (value materials in cash)?										
9. How many other jobs did (name) do during the last 12 months?										
10. How much in total has (name) earned during the last 12 months from these other jobs (value materials in cash)?										

Member ID	1	2	3	4	5	6	7	8	9	10
5. What are the problems you face with regard to your work? **1. Finance** **2. Difficult to access land** **3. Erratic weather conditions** **4. Uncertain demand for output** **5. Poor health** **6. Inadequate fish catch** **7. High cost of inputs e.g. fuel** **8. Difficulty in marketing products** **9. Unavailability of inputs** **10. Low price of products** **11. Other (Specify)**										
6. What is your employment status in you main job? **1. Self employed with employees** **2. Self employed without employees** **3. Unpaid family worker** **4. Casual worker** **5. Regular employee** **6. Domestic employee** **7. Apprentice** **8. Other (Specify)**										
7. For whom do you work in the main job? **1. Public** **2. Private formal** **3. Private informal** **4. Semi-public/parastatal** **5. NGO/International organization** **6. Other (Specify)**										
8. How much has (name) earned during the last 12 months from this economic activity (value materials in cash)?										
9. How many other jobs did (name) do during the last 12 months?										
10. How much in total has (name) earned during the last 12 months from these other jobs (value materials in cash)?										

Member ID	1	2	3	4	5	6	7	8	9	10
11. Where does (Name) usually do his/her work? 1. Office 2. Home 3. Factory 4. Workshop 5. Own land/farm 6. Other land/farm 7. River/Ocean 8. Hotel/Restaurant 9. Store/shop/table shop 10. Street at a fixed location 11. Street not at a fixed location 12. Lorry Park 13. Somebody's home or veranda 14. School 15. Hospital/Clinic 16. Other (specify)										
12. If you are not working now, for how long have you been without a job?										
13. If you are not working, how do you support yourself? **1. Own savings** **2. Pension** **3. Remittance from relative** **4. Borrow from family/friends** **5. Gifts** **6. Credit purchases** **7. Depend on spouse** **8. Depend on parents** **9. Other (Specify)**										

Member ID	1	2	3	4	5	6	7	8	9	10
11. Where does (Name) usually do his/her work? 1. Office 2. Home 3. Factory 4. Workshop 5. Own land/farm 6. Other land/farm 7. River/Ocean 8. Hotel/Restaurant 9. Store/shop/table shop 10. Street at a fixed location 11. Street not at a fixed location 12. Lorry Park 13. Somebody's home or veranda 14. School 15. Hospital/Clinic 16. Other (specify)										
12. If you are not working now, for how long have you been without a job?										
13. If you are not working, how do you support yourself? **1. Own savings** **2. Pension** **3. Remittance from relative** **4. Borrow from family/friends** **5. Gifts** **6. Credit purchases** **7. Depend on spouse** **8. Depend on parents** **9. Other (Specify)**										

PART E: EDUCATION

i. GENERAL EDUCATION –ANSWER FOR ALL HOUSEHOLD MEMBERS 3 YEARS AND OLDER

GRADE LEVEL CODES: 0… NONE; 01…PRE-SCHOOL; 11…P1; 12…P2; 13…P3; 14…P4; 15…P5; 16…P6; 17…JSS1; 18…JSS2; 19…JSS3; 20…M1; 21…M2; 22…M3; 23…M4; 24…SSS1; 25…SSS2; 26…SSS3; 27…S1; 28…S2; 29…S3; 30…S4; 31…S5; 32…OTHER (SPECIFY)

Member ID	1	2	3	4	5	6	7	8	9	10
1. ID of Person Interviewed										
2. Has (Name) ever attended school? 1. Yes 2. No >> PART E ii	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No
3. What was the highest grade successfully completed? **(See GRADE LEVEL CODES above)**										
4. What was the highest educational qualification attained? *Qualification Codes* 1. None 2. MSLC 3. BECE 4. Voc/Comm 5. Teacher Training 6. Teacher Post Secondary 7. GCE O Level 8. SSCE 9. GCE A Level 10. Tech/ Prof. Cert. 11. Tech/Prof. Dip. 12. HND 13. Bachelor's 14. Master's 15. Other (specify)	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15.
5 Did (Name) attend school / college at any time during the last 12 months? 1. Yes 2. No >> PART E ii	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No
6. Is (Name) still in school? 1. Yes 2. No>>11	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No
7. Is the school (Name) attending Public or Private? 1. Public Religious 2. Public Non-religious 3. Private Religious 4. Private Non-religious	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.
8. What is the current grade? See *Grade Level Codes* Above										

ii: Literacy And Apprenticeships –Answer For Household Members 5 Years And Older

Member ID	1	2	3	4	5	6	7	8	9	10
14. In what Ghanaian language can (Name) read a phrase/ sentence? (SHOW FLASH CARD) **State the one in which you are most proficient** (see language codes below)										
15. In what Ghanaian language can (Name) write a phrase/ sentence? **State the one in which you are most proficient** (see language codes below)										
16. Can (Name) read a phrase/ sentence in English? **(SHOW FLASH CARD)** 1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No
17. Can (Name) write a sentence in English? 1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No
18. What languages does (Name) speak most fluently? (list all - see Codebook)										
19. Can (Name) do written calculations? (USE FLASH CARDS) 1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No
20. Has (Name) ever attended a literacy course? 1. Yes >>22 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No
21. If not attending/ attended, why? **(>> 23)** 1. Low quality 2. Not available 3. Do not need 4. Too costly 5. Takes much time 6. Not useful 7. Too far 8. Spouse does not want 9. Other, specify	1. 2. 3. 4. 5. 6. 7. 8. 9.	1. 2. 3. 4. 5. 6. 7. 8. 9.	1. 2. 3. 4. 5. 6. 7. 8. 9.	1. 2. 3. 4. 5. 6. 7. 8. 9.	1. 2. 3. 4. 5. 6. 7. 8. 9.	1. 2. 3. 4. 5. 6. 7. 8. 9.	1. 2. 3. 4. 5. 6. 7. 8. 9.	1. 2. 3. 4. 5. 6. 7. 8. 9.	1. 2. 3. 4. 5. 6. 7. 8. 9.	1. 2. 3. 4. 5. 6. 7. 8. 9.

Member ID	1	2	3	4	5	6	7	8	9	10
9. How much travelling time does (Name) spend going to and from school daily? 1. Hours 2. Minutes (IF IN A BOARDING SCHOOL CODE: HRS …00; MINS … 00) >> 11	h m	h m	h m	h m	h m	h m	h m	h m	h m	h m
10. In the last 12 months, how much was spent on (Name) for transportation to and from school? 1. GHc 2. P	c p	c p	c p	c p	c p	c p	c p	c p	c p	c p
11. What was the total schooling expenditure spent on (Name) in the last 12 months? 1. GHc 2. P	c p	c p	c p	c p	c p	c p	c p	c p	c p	c p
12. Who paid for most of the educational expenses? 1. Father 2. Mother 3. Both parents 4. Other HH member 5. Other relative 6. Non-relative 7. Self 8. Other (specify)	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.
13. What mode of transportation does (Name) use to go to school? (If boarding, how does (Name) get there?) 1. Walk 2. Bicycle 3. Car 4. Bus 5. Train 6. Boat 7. Animal 8. Other (Specify)	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.	1. 2. 3. 4. 5. 6. 7. 8.

Member ID	1	2	3	4	5	6	7	8	9	10
22. For how many months has (Name) been attending / attended a literacy course? 1. months										
23. Is (Name) or has (Name) ever been an apprentice? 1. Yes, currently 2. Yes, in past 3. No >> 29	1. 2. 3.	1. 2. 3.	1. 2. 3	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3	1. 2. 3.	1. 2. 3.
24. How long is (was) the apprentice-ship? 1. months										
25. Describe the industry that this trade is connected to?										
26. Describe the sector that this trade is connected with. **(See Codebook for the Ghana informal sector codes)**										
27. Did (Name) pay a fee for this training? 1. Yes, in kind 2. Yes, in cash 3. Both 4. No >>29	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.
28. How much did (Name) pay for the training? 1. cedis 2. pesewas	c p	c p	c p	c p	c p	c p	c p	c p	c p	c p
29. Has (Name) ever attended other short training courses lasting not more than 6 months? 1. Yes 2. No >> PART G	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No	1. Yes 2. No
30. What was the main subject of the most recent training? 1. Clerical 2. Prof/Managerial 3. Computer 4. Marketing 5. Teaching 6. Leadership 7. Medicine 8. Accountancy 9. Trade/Skill 10. Other, specify	1. 2. 3. 4. 5. 6. 7. 8. 9. 10.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10.	1. 2. 3. 4. 5. 6. 7. 8. 9. 10.

Section 2:

Part A: HEALTH IN THE LAST 4 WEEKS

Member ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1a. During the past four weeks has [Name] suffered any illness or injury? 1. Neither >> 8 2. Illness 3. Injury >> 8 4. Both	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.	1. 2. 3. 4.
1b. When was the last time [Name] suffered from any illness or injury? 1. Within the last 2 weeks 2. Beyond 2 weeks																		
2. What was the illness that [Name] suffered? 1. Watery diarrhea 2. Diarrhea with blood >> 8 3. Fever >> 8 4. Cold/cough >> 8 5. Guinea Worm >> 8 6. Bilharzia >>F8 7. Other (specify) >> 8	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____	1. 2. 3. 4. 5. 6. 7. _____
3. What was the most important liquid that was given to [Name] to drink? 1. Oral Rehydration Salt (ORS) 2. Rice Water 3. Soup 4. Homemade sugar/saltwater solution, 5. Milk or infant formula, 6. Coconut water, 7. Water	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.	1. 2. 3. 4. 5. 6. 7.
4. If watery diarrhea, was [Name] given the same amount to drink as before the watery diarrhea, or more, or less? 1. Same 2. More 3. Less	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.
5. For how many days during the last 4 weeks has [Name] suffered from this condition? (1-14 days)																		

Member ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
12. What was the name of the health care facility? Enter name 1.Name																		
13. Code of Facility -																		
14. Is this a public or private facility? 1. Public, 2. Private, 3. Private non-religious	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.	1. 2. 3.
15. How much did [Name] pay for this consultation? (GH cedis and pesawas)	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p
16. How much did [Name] pay to travel and return? (GH cedis and pesawas)	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p
17. How much time did it take to travel to and from the facility? (TRAVEL TIME) 1. Hours 2. Minutes	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m
18. How long did [Name] wait in the health facility before being attended to by a health officer? (WAITING TIME) 1. Hours 2. Minutes	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m
19. How much TOTAL time (Waiting and Receiving treatment) did [Name] spend at the health facility? 1. Hours 2. Minutes	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m	___h ___m
20. During last 4 weeks was [Name] admitted to a hospital/ health care facility on account of an illness/ injury? (Include traditional healing centers) 1. Yes, 2. No >> 23	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N
21. How many nights did [Name] spend in the health center during the past 4 weeks? (1-14 nights)																		

Member ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
6. During the last 4 weeks did [Name] have to stop the usual activities because of this condition? 1. Yes, 2. No (>> 8)	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N
7. For how many days did [Name] have to stop his/her usual activity?(1-14 days)																		
8. During the last 4 weeks, has [Name] consulted any health care facility? 1. Yes 2. No >> 23	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N
9. On the most recent visit, who did [Name] consult? 1. Doctor, 2. Dentist, 3. Nurse, 4. Medical Asst., 5. Midwife, 6. Pharmacist, 7. Drug/Chemical seller, 8. Traditional Healer, 9. Trained TBA, 10. Untrained TBA, 11. Spiritualist, 12. Other (specify)	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 5 6 7 8 9 10 11 12
10. What was the reason for the most recent visit? 1. Illness, 2. Injury, 3. Follow-up, 4. Check-up, 5. Prenatal care, 6. Postnatal care, 7. Vaccination, 8. Other (specify)	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8
11. Where did the consultation take place? 1. Hospital, 2. Clinic, 3. MCH Clinic, 4. Maternity home, 5. Pharmacy, 6. Chemical Store, 7. Consultant's home, 8.Patient's home, 9. Community epicenter, 10. Other (specify)	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____	1. 2. 3. 4. 5. 6. 7. 8. 9. 10. _____

Member ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
22. How much did/will [Name] pay for staying in the health center during the past 4 weeks? (GH cedis and pesawas)	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p
23. During the last 4 weeks did [Name] purchase any medicine or medical supplies? 1. Yes, 2. No >> Part B	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N
24. How much did [Name] pay altogether for these medicine/medical supplies? (GH cedis and pesawas)	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p
25. Total medical expenses over the past 4 weeks (If cannot, give breakdown) (GH cedis and pesawas)	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p
26. Were all the medical/medicinal supplies obtainable from the health facility? 1. Yes, 2. No	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N

PART B: HEALTH IN LAST 12 MONTHS

Member ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. For the last 12 months, was [Name] hospitalized for any injury? 1. Yes, 2. No	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N	1. Y 2. N
2. If yes, how many days was NAME hospitalized? 1.number of days																		
3. What is the total cost of hospitalization? (GH cedis and pesawas)	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p
4. Who pays for the majority of medical expenses including	1. 2.	1. 2.	1. 2.	1. 2.	1. 2.	1. 2.	1. 2.	1. 2.	1. 2.	1. 2.	1. 2.	1. 2.	1. 2.	1. 2.	1. 2.	1. 2.	1. 2.	1. 2.

consultations and hospital stays (if any)? *ID of household member,* 1. Other relative 2. Government 3. Employer 4. Household member's employer 5. Health insurance 6. Other (specify)	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.	3. 4. 5. 6.
5. During the last 12 months did (NAME) suffer from onchocerciasis? 1. Yes 2. No>> 7	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2
6. What were the effects of onchocerciasis that name suffered from? Circle all that apply 1. Blindness 2 .Body itches 3 .Skin rashes 4 .Other (specify)	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
7. Did (NAME) take any oncho control drugs during the last 12 months? 1. Yes >> 9 2. No	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2
8. Why did (NAME) not take the oncho control drug 1. drug not available 2. because of side effect of drug previously 3.Drug not useful 4.did not trust distributer 5.Oother (specify) (>> 19)	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
9a. How many tablets did (NAME) take during the last 12 months?																			
9b. How many times has [Name] taken the oncho drugs during the last 12 months?																			
10. Did name experience any reactions from the oncho control drug that (NAME) took? 1. Yes 2. No>> 12	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2

11. If yes, list the reactions (NAME) experienced from the oncho control drug? 1.Nausea 2.Laziness 3.Sleepiness 4.Body pains 5.Other (specify)	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
12. Did (NAME) pay for the oncho control drug that (NAME) took 1. Yes 2.No >> 14	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2
13. If yes, how much did name pay for the oncho control drug that (NAME) took (GH cedis and pesawas)	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p	___c ___p
14. Did name incur any other cost on the treatment of oncho during the last 12 months? 1. Yes 2. No >> 17	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2
15. List those other means of controlling oncho on which (NAME) incurred cost.																			
16. Give the total of all the other cost incurred on the control of oncho during the last 12 months. (GH cedis and pesawas)	___c p	___c p	___c p	___c p	___c p	___c p	___c p	___c p	___c p	___c p	___c p	___c p	___c p	___c p	___c p	___c p	___c p	___c p	___c p
17. During the last 12 months did (NAME) have to stop his/her usual activity because of oncho? 1. Yes 2. No	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2	1 2

18. If yes to 17, how many days did (NAME) have to stop his/her usual activity because of oncho?																			
19. When was the first time [Name] took oncho drugs? Year [yyyy] [CODE 9999 if Never and move to next person]																			
20. How many times (years) has [Name] taken oncho drugs? Years																			
21. When was the last time [Name] took any oncho drugs? 1. **Year [yyyy]**																			

Section 3: Assets
Part A: Farm Assets
i. Animals in the Household

0. Does your household own livestock that it uses for production or for food?

1. Yes 2. No >> Part A ii.

	A. Draught animal (donkey, horse, bullock)	**B. Cattle**	**C. Sheep**	**D. Goats**	**E. Pigs**	**F. Rabbits**	**G. Chickens/ Roosters**	**H. Fish**	**I. Other 1**	**J. Other 2**
1. How many (…) are currently possessed by people in this household? 1. Number (if zero, next animal)	1. |_|_| no.	1. |_|_| no.	1. |_|_| no.	1. |_|_| no.	1. |_|_| no.	1. |_|_| no.	1. |_|_| no.	1. |_|_| no.	1. |_|_| no.	1. |_|_| no.
2. List the IDs of up to two people who own this animal? 1. Owner 1 2. Owner 2 3. Owned by whole household	1. ____ ID 2. ____ ID 3. ____HH	1. ____ ID 2. ____ ID 3. ____HH	1. ____ ID 2. ____ ID 3. ____HH	1. ____ ID 2. ____ ID 3. ____HH	1. ____ ID 2. ____ ID 3. ____HH	1. ____ ID 2. ____ ID 3. ____HH	1. ____ ID 2. ____ ID 3. ____HH	1. ____ ID 2. ____ ID 3. ____HH	1. ____ ID 2. ____ ID 3. ____HH	1. ____ ID 2. ____ ID 3. ____HH
3. What is the current value of these animals, if you sold all of them? 1. Cedis 2. Pesewas	c p	c p	c p	c p	c p	c p	c p	c p	c p	c p
4. In the last 12 months, have you spent money to maintain this animal? 1. Yes 2. No >> 6	1. 2. >>6	1. 2. >>6	1. 2. >>6	1. 2. >>6	1. 2. >>6	1. 2. >>6	1. 2. >>6	1. 2. >>6	1. 2. >>6	1. 2. >>6
5. How much have you spent on <u>fodder/ feed produced outside your home</u> in the last 12 months? 1. cedis 2. pesewas	c p	c p	c p	c p	c p	c p	c p	c p	c p	c p
6. What is the value of <u>fodder/feed produced within your home</u> in the last 12 months? 1. cedis 2. pesewas	c p	c p	c p	c p	c p	c p	c p	c p	c p	c p
7. How much has been spent on <u>veterinary expenses</u> for (…) in the last 12 months? 1. Cedis 2. Pesewas	c p	c p	c p	c p	c p	c p	c p	c p	c p	c p
8. How much has been spent on <u>enclosure, shelter, restraint and related expenses</u> for (…) in the last 12 months? 1. Cedis 2. Pesewas	c p	c p	c p	c p	c p	c p	c p	c p	c p	c p

	A. Draught animal (donkey, horse, bullock)	B. Cattle	C. Sheep	D. Goats	E. Pigs	F. Rabbits	G. Chickens/ Roosters	H. Fish _____	I. Other 1 _____	J. Other 2 _____
9. How much has been spent on hired labor to maintain (…) in the last 12 months? 1. Cedis 2. Pesewas	c p	c p	c p	c p	c p	c p	c p	c p	c p	c p
10. How much has been spent on access to water to maintain (…) in the last 12 months? 1. Cedis 2. Pesewas	c p	c p	c p	c p	c p	c p	c p	c p	c p	c p
11. How much has been spent on other expenses to maintain (…) in the last 12 months? 1. Cedis 2. Pesewas	c p	c p	c p	c p	c p	c p	c p	c p	c p	c p
12. How much revenue have you earned from the sales of (…) and (…) products in the last 12 months?	c p >> Next animal	c p >> Next animal	c p >> Next animal	c p >> Next animal	c p >> Next animal	C p >> Next animal	c p >> Next animal	c p >> Next animal	c p >> Next animal	c p >> Next Section

	A. Number	B. ID of Two Owners	C. Current Value
21. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p
22. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p
23. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p
24. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p
25. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p
26. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p
27. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p
28. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p
29. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p
30. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p

ii. Tools
13.
a.) How many of the following tools (in working condition) does your household own?
b.) List two owners of these tools. If all are owned by whole household, code HH
c.) What is the current value this tool/these tools?

	A. Number	B. ID of Two Owners	C. Current Value
1. 2-wheel tractor	a. _____	b. _____ ID _____ ID	c. _____c _____ p
2. 4-wheel tractor	a. _____	b. _____ ID _____ ID	c. _____c _____ p
3. Cane crusher	a. _____	b. _____ ID _____ ID	c. _____c _____ p
4. Hoe	a. _____	b. _____ ID _____ ID	c. _____c _____ p
5. Axe	a. _____	b. _____ ID _____ ID	c. _____c _____ p
6. Rake	a. _____	b. _____ ID _____ ID	c. _____c _____ p
7. Shovel	a. _____	b. _____ ID _____ ID	c. _____c _____ p
8. Pick(-axe)	a. _____	b. _____ ID _____ ID	c. _____c _____ p
9. Sickle/reaping hook	a. _____	b. _____ ID _____ ID	c. _____c _____ p
10. Harrow	a. _____	b. _____ ID _____ ID	c. _____c _____ p
11. Tiller	a. _____	b. _____ ID _____ ID	c. _____c _____ p
12. Water pump	a. _____	b. _____ ID _____ ID	c. _____c _____ p
13. Irrigation pipe	a. _____	b. _____ ID _____ ID	c. _____c _____ p
14. Cutlass	a. _____	b. _____ ID _____ ID	c. _____c _____ p
15. Trailer / cart	a. _____	b. _____ ID _____ ID	c. _____c _____ p
16. Spraying machine	a. _____	b. _____ ID _____ ID	c. _____c _____ p
17. Outboard motor	a. _____	b. _____ ID _____ ID	c. _____c _____ p
18. Canoe	a. _____	b. _____ ID _____ ID	c. _____c _____ p
19. Plough	a. _____	b. _____ ID _____ ID	c. _____c _____ p
20. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p

iii. Durable goods
14.
a.) How many of the following goods (in working condition) does your household own?
b.) List two owners of these good. if owned by whole household, code HH
c.) What is the current value of each good on average?

	A. Number	B. ID of Two owners	C. Current Value
1. Room Furniture	a. _____	b. _____ ID _____ ID	c. _____c _____ p
2. Sewing Machine	a. _____	b. _____ ID _____ ID	c. _____c _____ p
3. Stove (kerosene)	a. _____	b. _____ ID _____ ID	c. _____c _____ p
4. Stove (electric)	a. _____	b. _____ ID _____ ID	c. _____c _____ p
5. Stove (gas)	a. _____	b. _____ ID _____ ID	c. _____c _____ p
6. Refrigerator	a. _____	b. _____ ID _____ ID	c. _____c _____ p
7. Freezer	a. _____	b. _____ ID _____ ID	c. _____c _____ p
8. Air conditioner	a. _____	b. _____ ID _____ ID	c. _____c _____ p
9. Fan	a. _____	b. _____ ID _____ ID	c. _____c _____ p
10. Radio	a. _____	b. _____ ID _____ ID	c. _____c _____ p
11. Radio Cassette Player	a. _____	b. _____ ID _____ ID	c. _____c _____ p
12. Record Player	a. _____	b. _____ ID _____ ID	c. _____c _____ p
13. 3-in-1 radio system	a. _____	b. _____ ID _____ ID	c. _____c _____ p
14. Video Player	a. _____	b. _____ ID _____ ID	c. _____c _____ p
15. Desktop Computer	a. _____	b. _____ ID _____ ID	c. _____c _____ p
16. Laptop Computer	a. _____	b. _____ ID _____ ID	c. _____c _____ p

	A. Number	B. ID of Two owners	C. Current Value
17. Printer	a. _____	b. _____ ID _____ ID	c. _____c _____ p
18. Computer accessories	a. _____	b. _____ ID _____ ID	c. _____c _____ p
19. camcorder/video camera	a. _____	b. _____ ID _____ ID	c. _____c _____ p
20. Camera	a. _____	b. _____ ID _____ ID	c. _____c _____ p
21. Iron (electric)	a. _____	b. _____ ID _____ ID	c. _____c _____ p
22. Bicycle	a. _____	b. _____ ID _____ ID	c. _____c _____ p
23. Motorcycle	a. _____	b. _____ ID _____ ID	c. _____c _____ p
24. Books (Textbooks)	a. _____	b. _____ ID _____ ID	c. _____c _____ p
25. Cell (mobile) Phone handset	a. _____	b. _____ ID _____ ID	c. _____c _____ p
26. Telephone (Fixed line)	a. _____	b. _____ ID _____ ID	c. _____c _____ p
27. TV	a. _____	b. _____ ID _____ ID	c. _____c _____ p
28. Parabolic satellite	a. _____	b. _____ ID _____ ID	c. _____c _____ p
29. Blenders	a. _____	b. _____ ID _____ ID	c. _____c _____ p
30. VCD/DVD Player	a. _____	b. _____ ID _____ ID	c. _____c _____ p
31. Washing machine/ driers	a. _____	b. _____ ID _____ ID	c. _____c _____ p
32. Lanterns/ gas lights	a. _____	b. _____ ID _____ ID	c. _____c _____ p
33. Buckets	a. _____	b. _____ ID _____ ID	c. _____c _____ p
34. Motor-driven lawn mower	a. _____	b. _____ ID _____ ID	c. _____c _____ p
35. Electric Power Generator	a. _____	b. _____ ID _____ ID	c. _____c _____ p

SECTION 4, Part A - HOUSING CHARACTERISTICS:

RENT

1 Do you pay rent for this dwelling?
Yes 1
No . . . . . 2 (=> Q.6)

2 How much does the household pay in cash towards the rent?
AMOUNT IN GH¢ & P: GH¢ P
TIME UNIT: SEE CODES

3 Does your household also supply goods or services in exchange for this dwelling?
Yes 1
No . . . . . 2 (=> Q.5)

4 What is the appropriate value of these goods and services provided by your household?
AMOUNT IN GH¢ & P: GH¢ P
TIME UNIT: SEE CODES

5 Is part or all of the rent paid by someone who is not a member of your household?
Yes, All 1
Yes, Part 2
No . . . . . 3 (=> 8)

6 Who pays?
Relative 1
Private employer 2
Government 3
Private individual / agency 4
Other (specify) 5

7 How much does this person pay?
AMOUNT IN GH¢ & P: GH¢ P
TIME UNIT: SEE CODES

8 How much did your household spend for construction or repair costs and painting in the past 12 months on this dwelling?
AMOUNT IN GH¢ & P: GH¢ P
TIME UNIT: SEE CODES

TIME UNIT

Daily 1
Weekly 2
Monthly 3
Quarterly 4
Half yearly 5
Yearly 6
Not applicable 0

	A. Number	B. ID of Two owners	C. Current Value
36. Car	a. _____	b. _____ ID _____ ID	c. _____c _____ p
37. Truck/ bus	a. _____	b. _____ ID _____ ID	c. _____c _____ p
38. Musical instruments	a. _____	b. _____ ID _____ ID	c. _____c _____ p
39. Game cards (chess, ludo, etc)	a. _____	b. _____ ID _____ ID	c. _____c _____ p
40. Toys	a. _____	b. _____ ID _____ ID	c. _____c _____ p
41. Cutlery/ Utensils	a. _____	b. _____ ID _____ ID	c. _____c _____ p
42. Cups and plates	a. _____	b. _____ ID _____ ID	c. _____c _____ p
43. Torches	a. _____	b. _____ ID _____ ID	c. _____c _____ p
44. Land (non-farm)	a. _____	b. _____ ID _____ ID	c. _____c _____ p
45. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p
46. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p
47. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p
48. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p
49. Other, specify	a. _____	b. _____ ID _____ ID	c. _____c _____ p

SECTION 4, Part A - HOUSING CHARACTERISTICS:

UTILITIES

9 What is the main source of water supply for this household?
DRINKING: GENERAL USE: SEE CODES

10 How far is this source of water from your dwelling?
DRINKING NUMBER DISTANCE UNIT
GENERAL USE: NUMBER DISTANCE UNIT

11 How long does it take to go to the main water source, get water and come back?
1. Number of minutes 2. Water on premises 3. Don't know

12 How regular is your **main** source of water supply?
NUMBER: TIME UNIT SEE CODES

13 How much water does your household use in a day?
QUANTITY:
UNITS: Litres 1
Gallons 2
Bucket (No. 34) 3

14 Do you treat your water in any way to make it safe to drink?
1. Yes 2. No =>16

15 What do you usually do to the water to make it safer to drink? **Record all items mentioned**
1. Boil 2. Add bleach / chlorine 3. Strain it through a cloth
4. Use a water filter (ceremic,sand,composite,etc) 5. Solar disinfection
6. Let it stand and settle 7. Other (specify) 8. Don't know

16 How is the **main** water supply system operated and managed?
Self 1
Community operated and managed 2
Community Water Sanitation Agency 3
Ghana Water Company Limited 4
NGO 5
Other (specify) 6
Not Applicable 7

17 Does the household pay a regular bill for this **main** water supply system?
Yes 1
No . . . . . 2 (=> 19)

CODES FOR Q. 9

Indoor plumbing 01
Inside standpipe 02
Water truck / tanker service 03
Water vendor 04
Pipe in neighbouring household 05
Private outside standpipe / tap 06
Public standpipe 07
Sachel / bottled water 08
Borehole 09
Protected well 10
Unprotected well 11
River / Stream 12
Rain water / spring 13
Dugout / pond / lake / dam 14
Other (specify) 15

DISTANCE CODES

In house 0
Yard 1
Metre 2
Kilometre 3
Mile 4

TIME UNIT

Daily 1
Weekly 2
Monthly 3
Quarterly 4
Half yearly 5
Yearly 6
Not applicable 0

SECTION 4, Part A - HOUSING CHARACTERISTICS I:

UTILITIES (cont'd.)

18 How much was your last bill? (only your part, if joint meter or shared bill).

AMOUNT IN GHq & P: GHq | P

TIME UNIT: SEE CODES

19 How much did your household pay to a private water vendor, neighbour or standpipe or any other source in the last 2 weeks?

AMOUNT IN GHq & P: GHq | P

20 Did your household sell any water to someone else?

Yes 1

No . . . . . 2 (>> Q.22)

21 How much did your household receive for the water sold in the last 2 weeks?

AMOUNT IN GHq & P: GHq | P

22 What is the main source of lighting for your dwelling?

Electricity (mains) 1
Kerosene . . . . . 2 (>> 24)
Gas lamp . . . . . 3 (>> 24)
Candles / Torches (flashlights) . . . . . 4 (>> 24)
Solar energy . . . . . 5 (>> 24)
Generator . . . . . 6 (>> 24)
No light . . . . . 7 (>> 24)
Other (specify) . . . . . 8 (>> 24)

23 How much was your last bill. (Only your part , if you have a joint meter or a shared bill)?

AMOUNT IN GHq & P: GHq | P

TIME UNIT: SEE CODES

24 What is the main fuel used by the household for cooking?

None, no cooking 1
Wood 2
Charcoal 3
Gas 4
Electricity 5
Kerosene 6
Crop residue / sawdust 7
Animal waste 8
Other (specify) 9

TIME UNIT	
Daily	1
Weekly	2
Monthly	3
Quarterly	4
Half yearly	5
Yearly	6
Not applicable . . .	0

SECTION 4, Part A - HOUSING CHARACTERISTICS I:

STRUCTURE AND AMENITIES

25 How does your household dispose of refuse?

Collected 1
Public dump 2
Dumped elsewhere . . . . . 3 (>> 27)
Burned by household . . . . . 4 (>> 27)
Buried by household . . . . . 5 (>> 27)
Other (specify) . . . . . 6 (>> 27)

26 How much does this household pay for refuse disposal?

AMOUNT IN GHq & P: GHq | P

TIME UNIT: SEE CODES

27 The last time (name of youngest child under 5 years) passed stools, what was done to dispose off the stools?

1. Child used toilet / latrine 2. Put / rinsed into drain or ditch
3. Thrown into garbage 4. Buried
5. Left it in the open 6. Other (specify)
7. Don't know

28 What type of toilet is used by your household?

Flush toilet . . . . . 1 (>> 30)
Pit latrine . . . . . 2 (>> 30)
KVIP . . . . . 3 (>> 30)
Pan / bucket . . . . . 4 (>> 30)
Public toilet (flush bucket / KVIP) 5
Toilet in another house . . . 6 (>> 30)
No toilet facility (bush, beach) . . . . . 7 (>> 30)
Other (specify) . . . . . 8 (>> 30)

29 How much does the household pay for the use of the toilet facility?

AMOUNT IN GHq & P: GHq | P

TIME UNIT: SEE CODES

30 What is the main construction material used for the outer wall of the main building?

Mud / Mud bricks 01
Wood / Bamboo 02
Metal sheet / Slate / Asbestos 03
Stone 04
Burned bricks 05
Cement / Sandcrete blocks 06
Landcrete 07
Thatch 08
Cardboard 09
Other (specify) 10

31 What is the main construction material used for the floor?

Earth / Mud / Mud bricks 1
Wood 2
Stone 3
Cement / Concrete 4
Burnt bricks 5
Vinyl tiles 6
Ceramic / Marble / Tiles 7
Terrazzo 8
Other (specify) 9

32 What is the main material used for the roof?

Palm leaves / Raffia / Thatch 1
Wood 2
Corrugated iron sheets 3
Cement / Concrete 4
Asbestos / Slate 5
Roofing tiles 6
Mud bricks / Earth 7
Bamboo 8
Other (specify) 9

33 Does the household have access to the following?

	Yes ..... 1 / No 2
Fixed Line Telephone	
Mobile phone	
Personal computer	
Internet	
E-commerce	
Pan Cattle Network (e.g M-NET)	

TIME UNIT	
Daily	1
Weekly	2
Monthly	3
Quarterly	4
Half yearly	5
Yearly	6
Not applicable	0

SECTION 4, Part B – HOUSING CHARACTERISTICS II.

To be asked to the household head. (Some of these questions can also be filled by interviewer).

1 In what type of dwelling does the household live?

Separate house (bungalow) 1
Semi-detached house 2
Flat / Apartment 3
Room(s) [Compound house] 4
Room(s) [Other type] 5
Several huts / buildings [same compound] 6
Several huts / buildings [different compounds] 7
Tents / improvised home 8
Other (specify) 9

2 How many rooms does this household occupy? (Count living rooms, dining rooms, bed rooms but not bathroom, toilet & kitchen).

NUMBER:

3 Do other households share this dwelling with you?

Yes 1

No 2

4 What is your present occupancy status?

Owning . . . . . 1 (>> 6)
Renting 2
Rent-free 3
Perching 4

5 Who owns this dwelling?

Relative not household member 1
Other private individual 2
Private employer 3
Other private agency 4
Public / Government ownership 5
Other (specify) 6

6 What is the condition of the house in which the household is living?

Good 1
Livable 2
Badly damaged 3

7 How are the surroundings of the house?

Clean 1
Average 2
Dirty 3

8 How many bedrooms does the household have?

NUMBER:

9 Is there a room used exclusively for cooking?

Yes 1

No . . . . . 2 (>> 12)

10 Is there a window in the room where cooking is done?

Yes 1

No 2

11 Is there a chimney / smoke outlet in the cooking place?

Yes 1

No 2

12 Is cooking done outside in any season (rainy season and dry season)?

Yes 1

No 2

13 If not all seasons, in which season is cooking done outside?

Rainy season 1

Dry season 2

14 Does your house have electricity?

Yes 1

No . . . . . 2 (>> 16)

SECTION 4, Part B – HOUSING CHARACTERISTICS II.

15 How regular is your power supply after sunset?

Regular 1
Cut once or twice a week 2
Cut more than twice a week 3

16 Distance to nearest public toilet in kilometres. (Code NA if there are no public toilets and people go to the bush, beach, etc.).

DISTANCE IN KM:

17 Average time spent travelling to and waiting at public toilet (in minutes).

TIME SPENT IN MINUTES:

18 Number of other buildings owned.

NUMBER:

19 Number of vacant plots owned.

NUMBER:

20 Where are animals kept at night?

Inside the house 1
Outside the house 2
Animal shed 3
NA . . . . . 4 (>> 23)

21 Where are animals tied during the day?

Inside the house 1
Outside the house 2
Animal shed 3
Free range 4

22 Where is hay (grass) kept?

Inside the house 1
Outside the house 2
Animal shed 3
No hay provided 4

23 Is there any open sewer / drain in / around the house?

Yes 1

No 2

Drains are covered 3

24 Is there garbage (trash) in / around the house?

Yes 1

No 2

Section 5a: Community Participation

Line household polling site (from village survey) _____

Ask for each household member.

HOUSEHOLD ID	1 List the social and religious organisations that you belong / participate in. CODES FROM VILLAGE ROSTER. SEPARATE WITH COMMAS.	2 How much does being a member of these groups benefit you individually? SEE CODES	3 What is the most important benefit that you feel you gain from being a member of these groups(s)? SEE CODES	4 Did you attend the previous village assembly meeting on [date]? Yes . . . *1 No . . . *2 (>> 7)	5 Did you speak about any issue at the meeting? Yes . . . *1 No . . . *2 (>> 7)	6 What was / were this issue(s)? SEE CODES. SEPARATE WITH COMMAS.	7 Did you vote in the previous election on [date]? Yes . . . *1 No . . . *2	8 How well do you think your current MP is doing? Excellent . . . *1 Very good . . . *2 Average . . . *3 Bad . . . *4 Very bad . . . *5	9 How well do you think your current assembly man is doing? Excellent . . . *1 Very good . . . *2 Average . . . *3 Bad . . . *4 Very bad . . . *5	CODES FOR Q. 2 Greatly . . . *1 Fairly . . . *2 A little . . . *3 Not at all . . . *4 (>> Next household member) CODES FOR Q. 3 Solidarity (common understanding . . . *1 Financial support . . . *2 Mutual education / learning . . . *3 Ability . . . *4 Public attributes . . . *5 Lobbying power . . . *6 Counselling . . . *7 Labour exchange . . . *8 Other (specify) . . . *9 CODES FOR Q. 6 Road or transport . . . *01 Health . . . *02 Education . . . *03 Irrigation or Water . . . *04 Other agricultural related . . . *05 Ceremonies . . . *06 Disputes . . . *07 Sanitation . . . *08 Construction other than roads or irrigation . . . *09 Access to and use of lands of the assembly . . . *10 Unemployment schemes . . . *11 Other (specify) . . . *12
01										
02										
03										
04										
05										
06										
07										
08										
09										
10										
11										
12										
13										
14										
15										
16										
17										
18										

Section 5b: Community Participation

1. How much do you trust each of the following types of people?	Not at all	Just a little	I trust them somewhat	I trust them a lot	Refused to answer	Do not know
A. Your relatives	1	2	3	4	77	99
B. Your neighbours	1	2	3	4	77	99
C. People with the same religion like yourself	1	2	3	4	77	99
D. People with a different religion	1	2	3	4	77	99
E. People from your tribe/ethnic group	1	2	3	4	77	99
F. People from different tribes/ethnic groups	1	2	3	4	77	99
G. People from your region of origin	1	2	3	4	77	99
H. People from different regions	1	2	3	4	77	99
I. People from other African countries	1	2	3	4	77	99

Appendix E. Literature review for morbidity and prevalence data

Study Site ID	Publication ID	Date entered	Curator	Study year	GBD estimate year	Bioclime	Season	Settlement Type	GBD Region	Country	Region / State	District / Ward / Municipality/River Basin	Location name	Latitude	Longitude	Elevation (m)	Primary vector	Secondary vector	Biting rate	Biting rate 95% CI	Population size	Sex	Lower age	Upper age	Diagnostic method	Incubation media	Incubation time	Endemicity level	Treatment rounds	Coverage est.	ss_exam	ss_pos	ss_prev	nod_exam	nod_pos	nod_prev	mf_measure	mf_unit	mf_denom	mf_exam	mf_prev
SOH1	4	2009.04.29	SOH	2004	2005	Forest	Rainy	Village	tin America, Centr	Mexico	Chiapas	Soconusco	Las Golondrinas	15.435	-92.654722	920	S. ochraceum s.l.		93815	91945 - 95723	303	Both	ALL		PCR	Saline	24 hours	Hyperendemic	29	90.1	285	7.98	2.8								
SOH2	4	2009.04.29	SOH	2006	2005	Forest	Rainy	Village	tin America, Centr	Mexico	Chiapas	Soconusco	Las Golondrinas	15.435	-92.654722	920	S. ochraceum s.l.		57857	51476 - 65007	317	Both	ALL		PCR	Saline	24 hours	Hyperendemic	37	90.1	313	0.939	0.3								
SOH3	4	2009.04.29	SOH	2004	2005	Forest	Rainy	Village	tin America, Centr	Mexico	Chiapas	Soconusco	Nubes II	15.3025	-92.473611	1080	S. ochraceum s.l.		26429	22921 - 30442	137	Both	ALL		PCR	Saline	24 hours	Hyperendemic	29	90.1	111	1.998	1.8								
SOH4	4	2009.04.29	SOH	2006	2005	Forest	Rainy	Village	tin America, Centr	Mexico	Chiapas	Soconusco	Nubes II	15.3025	-92.473611	1080	S. ochraceum s.l.		22667	20720 - 24782	118	Both	ALL		PCR	Saline	24 hours	Hyperendemic	37	90.1	105	0	0								
SOH5	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Unknown	Moca	Moca			1097	S. metallicum s.l.	S. ochraceum s.l.				Both	1	4	Skin snip	Unknown	Unknown	Hyperendemic			2	0	0								
SOH5	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Unknown	Moca	Moca			1097	S. metallicum s.l.	S. ochraceum s.l.				Both	5	9	Skin snip	Unknown	Unknown	Hyperendemic			11	3	27.272727								
SOH5	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Unknown	Moca	Moca			1097	S. metallicum s.l.	S. ochraceum s.l.				Both	10	19	Skin snip	Unknown	Unknown	Hyperendemic			24	11	45.833333								
SOH5	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Unknown	Moca	Moca			1097	S. metallicum s.l.	S. ochraceum s.l.				Both	20	29	Skin snip	Unknown	Unknown	Hyperendemic			18	17	94.444444								
SOH5	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Unknown	Moca	Moca			1097	S. metallicum s.l.	S. ochraceum s.l.				Both	30	39	Skin snip	Unknown	Unknown	Hyperendemic			14	11	78.57142 9								
SOH5	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Unknown	Moca	Moca			1097	S. metallicum s.l.	S. ochraceum s.l.				Both	40	49	Skin snip	Unknown	Unknown	Hyperendemic			15	11	73.333333								
SOH5	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Unknown	Moca	Moca			1097	S. metallicum s.l.	S. ochraceum s.l.				Both	50	+	Skin snip	Unknown	Unknown	Hyperendemic			18	16	88.888889								
SOH5	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Unknown	Moca	Moca			1097	S. metallicum s.l.	S. ochraceum s.l.				Both	Total	Total	Skin snip	Unknown	Unknown	Hyperendemic			102	69	67.647059			Arithmetic	mf/mg	Positives	69	4.9
SOH6	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Moca	Olas de Moca			762	S. metallicum s.l.	S. ochraceum s.l.				Both	10	19	Skin snip	Unknown	Unknown	Hyperendemic			9	7	77.777778								
SOH6	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Moca	Olas de Moca			762	S. metallicum s.l.	S. ochraceum s.l.				Both	20	29	Skin snip	Unknown	Unknown	Hyperendemic			14	12	85.714286								
SOH6	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Moca	Olas de Moca			762	S. metallicum s.l.	S. ochraceum s.l.				Both	30	39	Skin snip	Unknown	Unknown	Hyperendemic			6	6	100								
SOH6	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Moca	Olas de Moca			762	S. metallicum s.l.	S. ochraceum s.l.				Both	40	49	Skin snip	Unknown	Unknown	Hyperendemic			11	10	90.909091								
SOH6	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Moca	Olas de Moca			762	S. metallicum s.l.	S. ochraceum s.l.				Both	50	+	Skin snip	Unknown	Unknown	Hyperendemic			6	5	83.333333								
SOH6	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Moca	Olas de Moca			762	S. metallicum s.l.	S. ochraceum s.l.				Both	Total	Total	Skin snip	Unknown	Unknown	Hyperendemic			46	40	86.956522			Arithmetic	mf/mg	Positives	40	14.46
SOH7	415	2009.05.02	SOH	1965	1990	Forest	Dry	Group of villages	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Panama estates	14.388091	-91.372677	426 - 914	S. metallicum s.l.	S. ochraceum s.l.				Both	10	19	Skin snip	Unknown	Unknown	Mesoendemic			3	0	0								
SOH7	415	2009.05.02	SOH	1965	1990	Forest	Dry	Group of villages	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Panama estates	14.388091	-91.372677	426 - 914	S. metallicum s.l.	S. ochraceum s.l.				Both	20	29	Skin snip	Unknown	Unknown	Mesoendemic			1	0	0								
SOH7	415	2009.05.02	SOH	1965	1990	Forest	Dry	Group of villages	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Panama estates	14.388091	-91.372677	426 - 914	S. metallicum s.l.	S. ochraceum s.l.				Both	30	39	Skin snip	Unknown	Unknown	Mesoendemic			5	3	60								
SOH7	415	2009.05.02	SOH	1965	1990	Forest	Dry	Group of villages	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Panama estates	14.388091	-91.372677	426 - 914	S. metallicum s.l.	S. ochraceum s.l.				Both	40	49	Skin snip	Unknown	Unknown	Mesoendemic			2	1	50								
SOH7	415	2009.05.02	SOH	1965	1990	Forest	Dry	Group of villages	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Panama estates	14.388091	-91.372677	426 - 914	S. metallicum s.l.	S. ochraceum s.l.				Both	50	+	Skin snip	Unknown	Unknown	Mesoendemic			7	2	28.571429								
SOH7	415	2009.05.02	SOH	1965	1990	Forest	Dry	Group of villages	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Panama estates	14.388091	-91.372677	426 - 914	S. metallicum s.l.	S. ochraceum s.l.				Both	Total	Total	Skin snip	Unknown	Unknown	Mesoendemic			18	6	33.333333			Arithmetic	mf/mg	Positives	48	28.37
SOH8	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Yepocapa	14.388091	-91.372677	1250	S. metallicum s.l.	S. ochraceum s.l.				Both	10	19	Skin snip	Unknown	Unknown	Hyperendemic			4	4	100								
SOH8	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Yepocapa	14.388091	-91.372677	1250	S. metallicum s.l.	S. ochraceum s.l.				Both	20	29	Skin snip	Unknown	Unknown	Hyperendemic			6	6	100								
SOH8	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Yepocapa	14.388091	-91.372677	1250	S. metallicum s.l.	S. ochraceum s.l.				Both	30	39	Skin snip	Unknown	Unknown	Hyperendemic			5	3	60								
SOH8	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Yepocapa	14.388091	-91.372677	1250	S. metallicum s.l.	S. ochraceum s.l.				Both	40	49	Skin snip	Unknown	Unknown	Hyperendemic			4	3	75								
SOH8	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Yepocapa	14.388091	-91.372677	1250	S. metallicum s.l.	S. ochraceum s.l.				Both	50	+	Skin snip	Unknown	Unknown	Hyperendemic			14	9	64.285714								
SOH8	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Yepocapa	14.388091	-91.372677	1250	S. metallicum s.l.	S. ochraceum s.l.				Both	Total	Total	Skin snip	Unknown	Unknown	Hyperendemic			33	25	75.75756			Arithmetic	mf/mg	Positives	48	28.37
SOH9	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Panajabal	14.388091	-91.372677	686	S. metallicum s.l.	S. ochraceum s.l.				Both	1	9	Skin snip	Unknown	Unknown	Hyperendemic			1	1	100								
SOH9	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Panajabal	14.388091	-91.372677	686	S. metallicum s.l.	S. ochraceum s.l.				Both	10	19	Skin snip	Unknown	Unknown	Hyperendemic			6	6	100								
SOH9	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Panajabal	14.388091	-91.372677	686	S. metallicum s.l.	S. ochraceum s.l.				Both	20	29	Skin snip	Unknown	Unknown	Hyperendemic			4	4	100								
SOH9	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Panajabal	14.388091	-91.372677	686	S. metallicum s.l.	S. ochraceum s.l.				Both	30	39	Skin snip	Unknown	Unknown	Hyperendemic			3	3	100								
SOH9	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Panajabal	14.388091	-91.372677	686	S. metallicum s.l.	S. ochraceum s.l.				Both	40	49	Skin snip	Unknown	Unknown	Hyperendemic			3	3	100								
SOH9	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Panajabal	14.388091	-91.372677	686	S. metallicum s.l.	S. ochraceum s.l.				Both	50	+	Skin snip	Unknown	Unknown	Hyperendemic			7	7	100								
SOH9	415	2009.05.02	SOH	1965	1990	Forest	Dry	Village	tin America, Centr	Guatemala	Chimaltenango	Yepocapa	Panajabal	14.388091	-91.372677	686	S. metallicum s.l.	S. ochraceum s.l.				Both	Total	Total	Skin snip	Unknown	Unknown	Hyperendemic			24	24	100			Arithmetic	mf/mg	Positives	48	28.37
SOH10	415	2009.05.02	SOH	1965	1990	Forest	Dry	State	tin America, Centr	Guatemala	Chimaltenango					426 - 1250	S. metallicum s.l.	S. ochraceum s.l.				Both	1	4	Skin snip	Unknown	Unknown	Hyperendemic			2	0	0								
SOH10	415	2009.05.02	SOH	1965	1990	Forest	Dry	State	tin America, Centr	Guatemala	Chimaltenango					426 - 1250	S. metallicum s.l.	S. ochraceum s.l.				Both	5	9	Skin snip	Unknown	Unknown	Hyperendemic			12	4	33.333333								
SOH10	415	2009.05.02	SOH	1965	1990	Forest	Dry	State	tin America, Centr	Guatemala	Chimaltenango					426 - 1250	S. metallicum s.l.	S. ochraceum s.l.				Both	10	19	Skin snip	Unknown	Unknown	Hyperendemic			46	28	60.869565								
SOH10	415	2009.05.02	SOH	1965	1990	Forest	Dry	State	tin America, Centr	Guatemala	Chimaltenango					426 - 1250	S. metallicum s.l.	S. ochraceum s.l.				Both	20	29	Skin snip	Unknown	Unknown	Hyperendemic			41	39	95.097674								
SOH10	415	2009.05.02	SOH	1965	1990	Forest	Dry	State	tin America, Centr	Guatemala	Chimaltenango					426 - 1250	S. metallicum s.l.	S. ochraceum s.l.				Both	30	39	Skin snip	Unknown	Unknown	Hyperendemic			33	26	78.787879								
SOH10	415	2009.05.02	SOH	1965	1990	Forest	Dry	State	tin America, Centr	Guatemala	Chimaltenango					426 - 1250	S. metallicum s.l.	S. ochraceum s.l.				Both	40	49	Skin snip	Unknown	Unknown	Hyperendemic			35	28	80								
SOH10	415	2009.05.02	SOH	1965	1990	Forest	Dry	State	tin America, Centr	Guatemala	Chimaltenango					426 - 1250	S. metallicum s.l.	S. ochraceum s.l.				Both	50	+	Skin snip	Unknown	Unknown	Hyperendemic			52	39	75								
SOH10	415	2009.05.02	SOH	1965	1990	Forest	Dry	State	tin America, Centr	Guatemala	Chimaltenango					426 - 1250	S. metallicum s.l.	S. ochraceum s.l.				Both	Total	Total	Skin snip	Unknown	Unknown	Hyperendemic			223	164	73.542601	164	70	42.6829268					
SOH11	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Rio Macay	Rio Macay Upstream	2.933333	-77.2	50						Both	0	4	Skin snip	Saline	15mins	Mesoendemic			10	0	0								
SOH11	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Rio Macay	Rio Macay Upstream	2.933333	-77.2	50						Both	5	9	Skin snip	Saline	15mins	Mesoendemic			11	1.001	9.1								
SOH11	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Rio Macay	Rio Macay Upstream	2.933333	-77.2	50						Both	10	19	Skin snip	Saline	15mins	Mesoendemic			27	7.02	26								
SOH11	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Rio Macay	Rio Macay Upstream	2.933333	-77.2	50						Both	20	29	Skin snip	Saline	15mins	Mesoendemic			9	0.999	11.1								
SOH11	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Rio Macay	Rio Macay Upstream	2.933333	-77.2	50						Both	30	39	Skin snip	Saline	15mins	Mesoendemic			17	10.03	59								
SOH11	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Rio Macay	Rio Macay Upstream	2.933333	-77.2	50						Both	40	+	Skin snip	Saline	15mins	Mesoendemic			58	20.01	34.5								
SOH11	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Rio Macay	Rio Macay Upstream	2.933333	-77.2	50						Both	Total	Total	Skin snip	Saline	15mins	Mesoendemic			132	39.336	29.8								
SOH12	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Lopez	Lopez	2.825	-77.216	50						Both	0	4	Skin snip	Saline	15mins	Hyperendemic			0	0	0								
SOH12	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Lopez	Lopez	2.825	-77.216	50						Both	5	9	Skin snip	Saline	15mins	Hyperendemic			0	0	0								
SOH12	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Lopez	Lopez	2.825	-77.216	50						Both	10	19	Skin snip	Saline	15mins	Hyperendemic			41	0	0								
SOH12	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Lopez	Lopez	2.825	-77.216	50						Both	20	29	Skin snip	Saline	15mins	Hyperendemic			24	1.008	4.2								
SOH12	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Lopez	Lopez	2.825	-77.216	50						Both	30	39	Skin snip	Saline	15mins	Hyperendemic			10	0	0								
SOH12	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Lopez	Lopez	2.825	-77.216	50						Both	40	+	Skin snip	Saline	15mins	Hyperendemic			85	1.995	4.7								
SOH12	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Lopez	Lopez	2.825	-77.216	50						Both	Total	Total	Skin snip	Saline	15mins	Hyperendemic			150	4.65	3.1								
SOH13	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Rio Macay	Total	2.933333	-77.2	50						Both	0	4	Skin snip	Saline	15mins	Hyperendemic			10	0	0								
SOH13	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Rio Macay	Total	2.933333	-77.2	50						Both	5	9	Skin snip	Saline	15mins	Hyperendemic			11	1.001	9.1								
SOH13	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Rio Macay	Total	2.933333	-77.2	50						Both	10	19	Skin snip	Saline	15mins	Hyperendemic			68	7.004	10.3								
SOH13	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Rio Macay	Total	2.933333	-77.2	50						Both	20	29	Skin snip	Saline	15mins	Hyperendemic			33	2.013	6.1								
SOH13	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Rio Macay	Total	2.933333	-77.2	50						Both	30	39	Skin snip	Saline	15mins	Hyperendemic			27	10.017	37.1								
SOH13	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Rio Macay	Total	2.933333	-77.2	50						Both	40	+	Skin snip	Saline	15mins	Hyperendemic			143	24	16.78321 7								
SOH13	407	2009.05.02	SOH	Unknown	1990	Forest	Unknown	River basin	tin America, Centr	Colombia	Cauca	Rio Macay	Total	2.933333	-77.2	50						Both	Total	Total	Skin snip	Saline	15mins	Hyperendemic			292	44.092	15.1	292	20	6.84931507	Arithmetic	mf/ss	All	44	3.5
SOH14	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Northern Region	Karonga		-9.935035	33.934021							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			54	1	1.8518519								
SOH15	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Northern Region	Chitipa		-9.702289	33.270721							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			38	1	2.6315789								
SOH16	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Northern Region	Chelinda		-11.024102	33.857803							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			17	0	0								
SOH17	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Northern Region	Livingstonia		-10.601895	34.105854							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			46	0	0								
SOH18	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Northern Region	Rumphi		-10.882249	33.752866							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			27	0	0								
SOH19	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Northern Region	Mzuzu		-11.458491	34.019165							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			101	0	0								
SOH20	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Northern Region	Nkata Bay		-11.606501	34.264984							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			50	0	0								
SOH21	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Northern Region	Mzimba		-11.896916	33.596191							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			39	0	0								
SOH22	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Central Region	Nkhotakota		-12.969766	34.266357							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			318	0	0								
SOH23	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Central Region	Kasungu		-13.031325	33.484955							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			37	0	0								
SOH24	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Central Region	Ntchisi		-13.353045	33.921146							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			17	0	0								
SOH25	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Central Region	Dowa		-13.580544	34.000003							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			34	0	0								
SOH26	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Central Region	Salima		-13.783236	34.459476							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			170	0	0								
SOH27	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Central Region	Lilongwe		-13.996621	33.583338							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			228	2	0.877193								
SOH28	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Central Region	Mchinji		-13.747889	33.083332							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			17	0	0								
SOH29	402	2009.06.09	SOH	1969	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Southern Region	Cholo District		-16.06429	35.134277		S. woodi					Both	All	All	Skin snip	Saline	Unknown	Mesoendemic			1380	460	33.333333								
SOH30	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Southern Region	Chiromo		-16.551962	35.145264							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			49	1	2.0408163								
SOH31	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Southern Region	Nsanje		-16.748075	35.166668							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			202	0	0								
SOH32	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Southern Region	Fort Johnson	r called Mengochi dis	-14.486173	35.253304							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			117	0	0								
SOH33	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Southern Region	Ncheu	Now spelt Ntcheu	-14.749651	34.749756							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			67	1	1.4925373								
SOH34	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Southern Region	Zomba		-15.164585	35.500002							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			103	0	0								
SOH35	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Southern Region	Kasupe	istrict no longer exist	-15.172879	34.804688							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			70	0	0								
SOH36	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Southern Region	Mlanje	Now spelt Mulanje	-15.734101	35.583344		S. nyasaliendicum	S. cervicornatum				Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			208	1	0.5								
SOH37	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Southern Region	Chikwawa		-16.1646	34.750003							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			34	0	0								
SOH38	402	2009.06.09	SOH	1968	1990	Savannah	Unknown	District	r-Saharan Africa, E	Malawi	Southern Region	Chiradzulu		-15.697923	35.166668							Both	All	All	Skin snip	Saline	Unknown	Hyperendemic			67	0	0								
SOH39	402	2009.06.09	SOH	1969	1990	Savannah	Unknown	Village	r-Saharan Africa, E	Malawi	Southern Region	Cholo District	Byumbwe	-15.92	35.07		S. woodi					Both	All	All	Skin snip	Saline	Unknown	Mesoendemic			171	48	28.070175								
SOH40	402	2009.06.09	SOH	1969	1990	Savannah	Unknown	Village	r-Saharan Africa, E	Malawi	Southern Region	Cholo District	Kawombe	-16.06429	35.134277		S. woodi					Both	All	All	Skin snip	Saline	Unknown	Mesoendemic			120	42	35								
SOH41	402	2009.06.09	SOH	1969	1990	Savannah	Unknown	Village	r-Saharan Africa, E	Mal

Study Site ID	Publication ID	Date entered	Curator	Study year	GBD estimate year	Bioclime	Season	Settlement Type	GBD Region	Country	Region / State	Dsitrict / Ward / Municipality/River Basin	Location name	Latitude	Longitude	Elevation (m)	Primary vector	Secondary vector	Biting rate	Biting rate 95% CI	Population size	Sex	Lower age	Upper age	Diagnostic method	Incubation media	Incubation time	Endemicity level	Treatment rounds	Coverage est.	ss_exam	ss_pos	ss_prev	nod_exam	nod_pos	nod_prev	mf_measure	mf_unit	mf_denom	mf_exam	mf_prev
SOH66	360	2009.06.15	SOH	1970	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Biankouri	Nanergou	10.9	0.15		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				150	21	14								
SOH67	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Biankouri	Niali	10.883333	0.65		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				73	57.67	79								
SOH68	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Biankouri	Tami	10.85	0.016667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				134	80.4	60								
SOH69	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Ouke	Donga	10.833333	0.783333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				120	97.2	81								
SOH70	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Bamoina	Gale	10.833333	0.4		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				95	80.75	85								
SOH71	360	2009.06.15	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Bamoina	Gale	10.833333	0.4		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				416	240.032	57.7								
SOH72	360	2009.06.15	SOH	1970	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Bouankjaka	Touloung	10.8333	0.133333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				150	28.95	19.3								
SOH73	360	2009.06.15	SOH	1970	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Biankouri	Talongue	10.833333	0.033333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				150	39	26								
SOH74	360	2009.06.15	SOH	1970	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Biankouri	Nakre	10.8	0.066667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				150	31.95	21.3								
SOH75	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Biankouri	Wokambo	10.783333	-0.0333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				75	60.75	81								
SOH76	360	2009.06.15	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Biankouri	Wokambo	10.7833	-0.03333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				256	170.496	66.6								
SOH77	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Sansargou	Borgou	10.766667	0.566667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				159	138.33	87								
SOH78	360	2009.06.15	SOH	1976	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Sansargou	Borgou	10.7667	0.5667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				449	319.239	71.1								
SOH79	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Mandare	Sawaga	10.75	0.4		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				75	66	88								
SOH80	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Dapon	Bidjanga	10.716667	0.283333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				171	87.21	51								
SOH81	360	2009.06.15	SOH	1970	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango		Nano	10.7	0.1		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				150	22.05	14.7								
SOH82	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango		Nano	10.7	0.1		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				187	74.8	40								
SOH83	360	2009.06.15	SOH	1970	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango		Monne	10.7	0.083333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				150	25.95	17.3								
SOH84	360	2009.06.15	SOH	1970	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango		Geundoga	10.683333	0.2		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				150	22.05	14.7								
SOH85	360	2009.06.15	SOH	1970	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Koulougona	Nambiokou	10.683333	0.016667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				150	31.5	21								
SOH86	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango		Bogou	10.65	0.183333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				206	107.12	52								
SOH87	360	2009.06.15	SOH	1970	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango		Noklak	10.633333	0.2		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				150	28.05	18.7								
SOH88	360	2009.06.15	SOH	1970	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Yammboul	Bakoze	10.633333	0.066667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				150	33	22								
SOH89	360	2009.06.15	SOH	1970	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Yammboul	Niagou	10.616667	0.05		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				150	25.95	17.3								
SOH90	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Yammboul	Koundounguru	10.533333	0.15		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				128	93.44	73								
SOH91	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Dapango	Yammboul	Koumbone	10.416667	0.3		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				109	59.95	55								
SOH92	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Dapon	Kpinbonga	10.65	0.366667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				240	134.4	56								
SOH93	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Lakitaon	Loko	10.633333	0.25		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				386	289.5	75								
SOH94	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Koinnepouarbaga	Biaga	10.616667	0.45		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				113	57.63	51								
SOH95	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Sangabire	Glimba	10.533333	0.533333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				161	127.19	79								
SOH96	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Sangabire	Glimba	10.533333	0.533333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				59	41.89	71								
SOH97	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Laktaon	Galangachi	10.516667	0.35		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				241	149.42	62								
SOH98	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Laktaon	Galangachi	10.516667	0.35		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				133	86.45	65								
SOH99	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Tchanaga	10.5	0.5		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				139	112.59	81								
SOH100	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Tontondi	10.483333	0.55		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				243	184.68	76								
SOH101	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Nawakouou	10.466667	0.483333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				101	74.74	74								
SOH102	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Gbeti/Kousiogou	10.433333	0.6		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				106	75.26	71								
SOH103	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Mogou	10.383333	0.65		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				294	178.458	60.7								
SOH104	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Kounteore	10.366667	0.633333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				176	109.12	62								
SOH105	360	2009.06.15	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Gravillou	10.366667	0.483333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				95	70.3	74								
SOH106	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango		Diabouri	10.35	0.766667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				330	145.2	44								
SOH107	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Manian	10.35	0.433333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				207	144.9	70								
SOH108	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Manian	10.35	0.433333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				100	62	62								
SOH109	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango		Dje-dabou	10.333333	0.733333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				280	145.6	52								
SOH110	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango		Dje-dabou	10.333333	0.733333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				98	40.18	41								
SOH111	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Niambadé	10.3	0.416667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				33	21.78	66								
SOH112	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Sadori	10.283333	0.466667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				163	122.25	75								
SOH113	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Bountourougu	Koumongou	10.25	0.633333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				101	79.79	79								
SOH114	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Koumongokan	10.233333	0.4		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				150	106.5	71								
SOH115	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Koumongokan	10.233333	0.4		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				118	94.4	80								
SOH116	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Koumango	Nakpouekou	10.216667	0.483333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				155	91.45	59								
SOH117	360	2009.06.15	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Fare	10.183333	0.4		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				240	117.6	49								
SOH118	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Fare	10.183333	0.4		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				552	279.864	50.7								
SOH119	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Farego	10.15	0.383333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				122	21.96	18								
SOH120	360	2009.06.15	SOH	1976	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Oti	Pange	10.05	0.666667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				397	216.762	54.6								
SOH121	360	2009.06.15	SOH	1973	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango		Gongounou/Kango	10.033333	0.6		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				86	20.64	24								
SOH122	360	2009.06.15	SOH	1976	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Mango	Napoulok	Ounkpe	9.933333	0.65		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				102	67.32	66								
SOH123	360	2009.06.16	SOH	1976	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Pagouda	Kande	Keran	Tibra	10	1.116667		S. damnosum s.l.	S. squamosum s.s.			Both	All	All	Skin snip	Unknown	30mins				71	66.172	93.2								
SOH124	360	2009.06.16	SOH	1972	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Pagouda	Binah	Kongolo	9.9	1.333333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				190	49.4	26								
SOH125	360	2009.06.16	SOH	1972	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Pagouda	Binah	Bare	9.866667	1.116667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				226	122.04	54								
SOH126	360	2009.06.16	SOH	1972	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Pagouda	Binah	Konfess	9.8	1.233333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				306	140.76	46								
SOH127	360	2009.06.16	SOH	1972	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Pagouda	Binah	Panahore	9.783333	1.35		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				248	166.16	67								
SOH128	360	2009.06.16	SOH	1972	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Niamtougou	Oumou	Pouda	9.9	1.25		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				410	311.6	76								
SOH129	360	2009.06.16	SOH	1972	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Niamtougou	Poutouan	Tchore	9.85	1.016667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				240	165.6	69								
SOH130	360	2009.06.16	SOH	1976	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Niamtougou	Poutouan	Anima	9.766667	1.016667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				487	330.186	67.8								
SOH131	360	2009.06.16	SOH	1972	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Niamtougou	Kpelou	Yaka	9.733333	1.1		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				312	212.16	68								
SOH132	360	2009.06.16	SOH	1972	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Niamtougou	La Kara	Leon	9.7	0.95		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				200	176	88								
SOH133	360	2009.06.16	SOH	1976	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Niamtougou	La Kara	Leon	9.7	0.95		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				299	252.057	84.3								
SOH134	360	2009.06.16	SOH	1972	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Lama Kara	Kpedou	Kpatayou	9.7	1.166667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				190	104.5	55								
SOH135	360	2009.06.16	SOH	1972	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Lama Kara	Sossoa	Kasse	9.666667	1.266667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				312	193.44	62								
SOH136	360	2009.06.16	SOH	1976	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Lama Kara	La Kara	Kpesside	9.633333	0.95		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				227	182.962	80.6								
SOH137	360	2009.06.16	SOH	1972	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Lama Kara	La Kara	Sara Kawa	9.616667	1		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				302	217.44	72								
SOH138	360	2009.06.16	SOH	1972	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Lama Kara	La Kara	Landa-Pozenda	9.516667	1.283333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				327	264.87	81								
SOH139	360	2009.06.16	SOH	1976	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Lama Kara	La Kara	Landa-Pozenda	9.516667	1.283333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				432	345.168	79.9								
SOH140	360	2009.06.16	SOH	1972	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Lama Kara	La Kara	Bounon	9.466667	1.1		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				280	198.8	71								
SOH141	360	2009.06.16	SOH	1975	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Lama Kara	Agougou	Kolo	9.366667	1.316667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				115	37.95	33								
SOH142	360	2009.06.16	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bafilo		Soreda	9.316667	1.05		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				264	91.192	35.3								
SOH143	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari	La Kara	Katchamba	9.883333	0.533333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				96	24.96	26								
SOH144	360	2009.06.16	SOH	1975	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari	La Kara	Katchamba	9.883333	0.533333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				101	44.036	43.6								
SOH145	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari	La Kara	Tema	9.816667	0.516667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				99	0	0								
SOH146	360	2009.06.16	SOH	1975	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari	La Kara	Napimbo	9.816667	0.65		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				176	108.944	61.9								
SOH147	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari	Niangboa	Kadjol	9.8	0.966667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				160	72	45								
SOH148	360	2009.06.16	SOH	1975	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari	La Kara	Natchibore	9.8	0.75		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				121	87.967	72.7								
SOH149	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari		Kidjaboun	9.783333	0.466667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				135	20.25	15								
SOH150	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari		Gbangbale	9.683333	0.833333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				111	8.88	8								
SOH151	360	2009.06.16	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari	Tianko	Natchitipi	9.683333	0.733333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				779	374.699	48.1								
SOH152	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari	Tianko	Koutchicheou	9.6	0.766667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				187	41.14	22								
SOH153	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari	Oti	Nandouta	9.6	0.416667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				143	41.47	29								
SOH154	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari		Loajoule	9.366667	0.583333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				86	21.5	25								
SOH155	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari	Kama	Nangbani	9.266667	0.8		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				105	34.65	33								
SOH156	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari	Kama	Bigabo	9.216667	0.85		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				79	23.7	30								
SOH157	360	2009.06.16	SOH	1976	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari	Kama	Bigabo	9.216667	0.85		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				238	156.366	65.7								
SOH158	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari	Katcha	Dimouri	9.2	0.583333		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				20	11	55								
SOH159	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Bassari	Mo	Bangan	9.066667	0.7		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				105	80.85	77								
SOH160	360	2009.06.16	SOH	1976	1990	Savannah

Study Site ID	Publication ID	Date entered	Curator	Study year	GBD estimate year	Bioclime	Season	Settlement Type	GBD Region	Country	Region / State	Oistrict / Ward / Municipality/River Basin	Location name	Latitude	Longitude	Elevation (m)	Primary vector	Secondary vector	Biting rate	Biting rate 95% CI	Population size	Sex	Lower age	Upper age	Diagnostic method	Incubation media	Incubation time	Endemicity level	Treatment rounds	Coverage est.	ss_exam	ss_pos	ss_prev	nod_exam	nod_pos	nod_prev	mf_measure	mf_unit	mf_denom	mf_exam	mf_prev
SOH195	360	2009.06.16	SOH	1975	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Tabligbo	Haho	Esse-Zogbedji	6.583333	1.333333		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				422	132.086	31.3								
SOH196	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Tabligbo	Haho	Ahepe Agbleta	6.6	1.416667		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				140	26.6	19								
SOH197	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Tabligbo	Mono	Kini-Kondji	6.6	1.55		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				161	9.66	6								
SOH198	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Tabligbo	Mono	Tokpli	6.666667	1.6		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				96	39.36	41								
SOH199	360	2009.06.16	SOH	1975	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Tabligbo	Mono	Tokpli	6.666667	1.6		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				404	129.28	32								
SOH200	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Tabligbo	Mono	Gboto Vodougbe	6.666667	1.533333		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				104	31.2	30								
SOH201	360	2009.06.16	SOH	1975	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Tabligbo	Mono	Gboto Vodougbe	6.666667	1.533333		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				740	316.72	42.8								
SOH202	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Tabligbo	Mono	Togodo	6.733333	1.533333		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				56	22.96	41								
SOH203	360	2009.06.16	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Tabligbo	Haho	Yoto-Kope	6.7	1.316667		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				79	60.83	77								
SOH204	360	2009.06.16	SOH	1975	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Nuatja	Mono	Sika-Kondji	6.633333	1.583333		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				240	80.64	33.6								
SOH205	360	2009.06.16	SOH	1975	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Nuatja	Mono	Esse-Ana	6.716667	1.566667		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				514	266.766	51.9								
SOH206	360	2009.06.16	SOH	1975	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Nuatja	Mono	Kouve	6.666667	1.416667		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				273	81.354	29.8								
SOH207	360	2009.06.16	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Nuatja	Mono	Tetetou	7.016667	1.5		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				238	194.446	81.7								
SOH208	360	2009.06.16	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Nuatja	Kra	Kokpli	7.15	1.283333		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				244	185.44	76								
SOH209	360	2009.06.16	SOH	1976	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Nuatja	Mono	Ciame	7.266667	1.5		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				146	122.932	84.2								
SOH210	360	2009.06.16	SOH	1976	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Nuatja	Mono	Oualime	7.2	1.533333		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				136	93.024	68.4								
SOH211	360	2009.06.16	SOH	1976	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Nuatja	Mono	Ahassome	7.233333	1.5		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				96	74.976	78.1								
SOH212	360	2009.06.16	SOH	1976	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Nuatja	Mono	Petchikoe	7.166667	1.516667		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				96	81.024	84.4								
SOH213	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Amou	Zolou	7.3	1.15		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				91	56.42	62								
SOH214	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Mono	Atome	7.366667	1.533333		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				84	43.68	52								
SOH215	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Amou	Kpakpo	7.416667	1.25		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				96	41.28	43								
SOH216	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Mono	Akpare	7.466667	1.333333		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				103	37.08	36								
SOH217	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Kouffo	Agouna Deve	7.516667	1.633333		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				35	16.1	46								
SOH218	360	2009.06.24	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Mono	Onia Kope	7.516667	1.333333		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				141	116.466	82.6								
SOH219	360	2009.06.24	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Anie	Illougba	7.8	1.15		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				106	85.436	80.6								
SOH220	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Mono	Foukote	7.6	1.316667		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				82	29.52	36								
SOH221	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Anie	Awagorme	7.633333	1.2		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				119	67.83	57								
SOH222	360	2009.06.24	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Anie	Alamasou	7.683333	1.25		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				220	188.98	85.9								
SOH223	360	2009.06.24	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Mono	Kokote	8.016667	1.233333		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				135	105.165	77.9								
SOH224	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Mono	Kolokope	7.8	1.3		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				71	41.89	59								
SOH225	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Noukoue	Patalu	7.783333	1.533333		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				108	60.48	56								
SOH226	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Anie	Akaba Gare	7.95	1.05		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				72	43.92	61								
SOH227	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Ofe	Yebou Yebou	8.05	1.383333		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				109	69.76	64								
SOH228	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Ogou	Kamina	8.1	1.566667		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				130	84.5	65								
SOH229	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Atakpame	Mono	Isuati	8.45	1.366667		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				116	64.96	56								
SOH230	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Akposso and Badou	Amou	Tuchalpali	7.533333	1.116667		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				119	41.65	35								
SOH231	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Akposso and Badou	Amou	Adjahoun	7.483333	0.916667		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				121	15.73	13								
SOH232	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Akposso and Badou	Mamou	Agbo Kope	7.433333	0.666667		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				146	68.62	47								
SOH233	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Akposso and Badou	Gdadi	Agadja	7.5	0.783333		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				123	49.2	40								
SOH234	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Akposso and Badou	Domi	Betel	7.516667	0.583333		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				160	91.2	57								
SOH235	360	2009.06.24	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Akposso and Badou	Domi	Obe	7.533333	0.6		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				311	192.509	61.9								
SOH236	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Akposso and Badou		Betee	7.65	1.1		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				89	56.07	63								
SOH237	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Akposso and Badou		Mouna	7.633333	0.933333		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	30mins				104	66.56	64								
SOH238	360	2009.06.24	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Akposso and Badou	Ove	Zoubega Oga	7.566667	0.666667		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				177	109.209	61.7								
SOH239	360	2009.06.24	SOH	1974	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Akposso and Badou	Oua Oua	Klabe Soto	7.6	0.883333		S. squamosum s.s.					Both	All	All	Skin snip	Unknown	30mins				120	80.4	67								
SOH240	360	2009.06.24	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Soutoubousa	Anie	Fazao	8.7	0.766667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				561	334.917	59.7								
SOH241	360	2009.06.24	SOH	1977	1990	Savannah	Unknown	Village	-Saharan Africa, W	Togo	Tchamba	Mono	Diome	8.6	1.316667		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Unknown	30mins				295	226.855	76.9								
SOH242	355	2009.06.24	SOH	Unknown	1990	Savannah	Unknown	District	-Saharan Africa, W	Mali	Sikasso		Fourou	10.742837	-6.147881		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Water	30mins	Mesoendemic			269	100	37.174721								
SOH243	355	2009.06.24	SOH	Unknown	1990	Savannah	Unknown	District	-Saharan Africa, W	Mali	Bamako		Kolokani	13.570741	-8.041992		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Water	30mins	Mesoendemic			388	178	45.876289								
SOH244	355	2009.06.24	SOH	Unknown	1990	Savannah	Unknown	District	-Saharan Africa, W	Mali	Bamako		Didieni	13.883079	-8.094521		S. damnosum s.l.	S. squamosum s.s.				Both	All	All	Skin snip	Water	30mins	Hyperendemic			445	108	24.269663								
SOH246	355.2	2009.06.24	SOH	Unknown	1990	Savannah	Unknown	Village	-Saharan Africa, W	Burkina Faso	Upper Volta	Po District	Torem	11.167224	-1.142514		S. damnosum s.l.					Both	All	All	Skin snip	Water	30mins	Hyperendemic			205	129	62.926829								
SOH247	350	2009.06.25	SOH	1970	1990	Savannah	Average	Village	-Saharan Africa, W	Nigeria	Kainji Lake	Niger river	Komi	9.87957	4.614258		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	Unknown				56	26.992	48.2	56	9.016	16.1					
SOH248	350	2009.06.25	SOH	1970	1990	Savannah	Average	Village	-Saharan Africa, W	Nigeria	Kainji Lake	Niger river	Kpatachi	9.816674	4.633398		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	Unknown				198	104.94	53	198	62.964	31.8					
SOH249	350	2009.06.25	SOH	1970	1990	Savannah	Average	Village	-Saharan Africa, W	Nigeria	Kainji Lake	Niger river	Faku	9.844878	4.64241		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	Unknown				107	49.006	45.8	107	2.996	2.8					
SOH250	350	2009.06.25	SOH	1970	1990	Savannah	Average	Village	-Saharan Africa, W	Nigeria	Kainji Lake	Niger river	Awuru	9.700026	4.633355		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	Unknown				79	13.035	16.5	79	1.975	2.5					
SOH251	350	2009.06.25	SOH	1970	1990	Savannah	Average	Village	-Saharan Africa, W	Nigeria	Kainji Lake	Kainji Lake	Kuruwasa	10.397525	4.553833		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	Unknown				207	54.027	26.1	207	13.041	6.3					
SOH252	350	2009.06.25	SOH	1970	1990	Savannah	Average	Village	-Saharan Africa, W	Nigeria	Kainji Lake	Kainji Lake	Anfani	10.348695	4.550743		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	Unknown				290	38.86	13.4	290	31.03	10.7					
SOH253	350	2009.06.25	SOH	1970	1990	Savannah	Average	Village	-Saharan Africa, W	Nigeria	Kainji Lake	Oli river	Labarau	9.878064	4.546216		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	Unknown				17	12.002	70.6	17	2.006	11.8					
SOH254	350	2009.06.25	SOH	1970	1990	Savannah	Average	Group of villages	-Saharan Africa, W	Nigeria	Kainji Lake		Total	10.348695	4.550743		S. damnosum s.l.					Both	All	All	Skin snip	Unknown	Unknown				954	298.602	31.3	954	123.066	12.9					
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Male	0	4	Hoth 1.5mm	Other medium	24 hr				2	0	0								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Female	0	4	Hoth 1.5mm	Other medium	24 hr				2	0	0								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Male	5	9	Hoth 1.5mm	Other medium	24 hr				37	22	59.459459								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Female	5	9	Hoth 1.5mm	Other medium	24 hr				36	24	66.666667								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Male	10	14	Hoth 1.5mm	Other medium	24 hr				47	41	87.234043								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Female	10	14	Hoth 1.5mm	Other medium	24 hr				36	30	83.333333								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Male	15	19	Hoth 1.5mm	Other medium	24 hr				23	21	91.304348								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Female	15	19	Hoth 1.5mm	Other medium	24 hr				11	7	63.636364								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Male	20	24	Hoth 1.5mm	Other medium	24 hr				18	15	83.333333								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Female	20	24	Hoth 1.5mm	Other medium	24 hr				10	7	70								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Male	25	29	Hoth 1.5mm	Other medium	24 hr				21	19	90.47619								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Female	25	29	Hoth 1.5mm	Other medium	24 hr				15	14	93.333333								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Male	30	34	Hoth 1.5mm	Other medium	24 hr				30	29	96.666667								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Female	30	34	Hoth 1.5mm	Other medium	24 hr				16	13	81.25								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Male	35	39	Hoth 1.5mm	Other medium	24 hr				18	15	83.333333								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Female	35	39	Hoth 1.5mm	Other medium	24 hr				6	6	100								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Male	40	44	Hoth 1.5mm	Other medium	24 hr				16	15	93.75								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Female	40	44	Hoth 1.5mm	Other medium	24 hr				4	4	100								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Male	45	49	Hoth 1.5mm	Other medium	24 hr				11	9	81.818182								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Female	45	49	Hoth 1.5mm	Other medium	24 hr				1	1	100								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Male	50	+	Hoth 1.5mm	Other medium	24 hr				6	6	100								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Female	50	+	Hoth 1.5mm	Other medium	24 hr				7	7	100								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Male	All	All	Hoth 1.5mm	Other medium	24 hr				229	192	83.842795								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Female	All	All	Hoth 1.5mm	Other medium	24 hr				144	113	78.472222								
SOH255	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 19	6.267011	-10.342426		S. yahense	S. soubrense				Both	All	All	Hoth 1.5mm	Other medium	24 hr				746	610	81.769437								
SOH256	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 22	6.267011	-10.342426		S. yahense	S. soubrense				Male	0	4	Hoth 1.5mm	Other medium	24 hr				0	0	0								
SOH256	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 22	6.267011	-10.342426		S. yahense	S. soubrense				Female	0	4	Hoth 1.5mm	Other medium	24 hr				0	0	0								
SOH256	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 22	6.267011	-10.342426		S. yahense	S. soubrense				Male	5	9	Hoth 1.5mm	Other medium	24 hr				9	4	44.444444								
SOH256	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 22	6.267011	-10.342426		S. yahense	S. soubrense				Female	5	9	Hoth 1.5mm	Other medium	24 hr				0	0	0								
SOH256	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 22	6.267011	-10.342426		S. yahense	S. soubrense				Male	10	14	Hoth 1.5mm	Other medium	24 hr				12	6	50								
SOH256	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 22	6.267011	-10.342426		S. yahense	S. soubrense				Female	10	14	Hoth 1.5mm	Other medium	24 hr				5	4	80								
SOH256	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 22	6.267011	-10.342426		S. yahense	S. soubrense				Male	15	19	Hoth 1.5mm	Other medium	24 hr				11	9	81.818182								
SOH256	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 22	6.267011	-10.342426		S. yahense	S. soubrense				Female	15	19	Hoth 1.5mm	Other medium	24 hr				5	5	100								
SOH256	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 22	6.267011	-10.342426		S. yahense	S. soubrense				Male	20	24	Hoth 1.5mm	Other medium	24 hr				7	6	85.714286								
SOH256	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 22	6.267011	-10.342426		S. yahense	S. soubrense				Female	20	24	Hoth 1.5mm	Other medium	24 hr				3	3	100								
SOH256	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 22	6.267011	-10.342426		S. yahense	S. soubrense				Male	25	29	Hoth 1.5mm	Other medium	24 hr				7	6	85.714286								
S

Study Site ID	Publication ID	Date entered	Curator	Study year	GBD estimate year	Bioclime	Season	Settlement Type	GBD Region	Country	Region / State	Dsitrict / Ward / Municipality/River Basin	Location name	Latitude	Longitude	Elevation (m)	Primary vector	Secondary vector	Biting rate	Biting rate 95% CI	Population size	Sex	Lower age	Upper age	Diagnostic method	Incubation media	Incubation time	Endemicity level	Treatment rounds	Coverage est.	ss_exam	ss_pos	ss_prev	nod_exam	nod_pos	nod_prev	mf_measure	mf_unit	mf_denom	mf_exam	mf_prev
SOH257	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 23	6.267011	-10.342426		S. yahense	S. soubrense				Male	50	+	Hoth 1.5mm	Other medium	24 hr			3	3	100								
SOH257	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 23	6.267011	-10.342426		S. yahense	S. soubrense				Female	50	+	Hoth 1.5mm	Other medium	24 hr			1	1	100								
SOH257	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 23	6.267011	-10.342426		S. yahense	S. soubrense				Male	All	All	Hoth 1.5mm	Other medium	24 hr			53	45	84.90566								
SOH257	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 23	6.267011	-10.342426		S. yahense	S. soubrense				Female	All	All	Hoth 1.5mm	Other medium	24 hr			20	15	75								
SOH257	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 23	6.267011	-10.342426		S. yahense	S. soubrense				Both	All	All	Hoth 1.5mm	Other medium	24 hr			146	120	82.191781								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Male	0	4	Hoth 1.5mm	Other medium	24 hr			0	0	0								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Female	0	4	Hoth 1.5mm	Other medium	24 hr			0	0	0								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Male	5	9	Hoth 1.5mm	Other medium	24 hr			2	2	100								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Female	5	9	Hoth 1.5mm	Other medium	24 hr			3	0	0								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Male	10	14	Hoth 1.5mm	Other medium	24 hr			9	6	66.666667								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Female	10	14	Hoth 1.5mm	Other medium	24 hr			0	0	0								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Male	15	19	Hoth 1.5mm	Other medium	24 hr			1	1	100								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Female	15	19	Hoth 1.5mm	Other medium	24 hr			0	0	0								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Male	20	24	Hoth 1.5mm	Other medium	24 hr			3	2	66.666667								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Female	20	24	Hoth 1.5mm	Other medium	24 hr			1	1	100								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Male	25	29	Hoth 1.5mm	Other medium	24 hr			0	0	0								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Female	25	29	Hoth 1.5mm	Other medium	24 hr			0	0	0								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Male	30	34	Hoth 1.5mm	Other medium	24 hr			0	0	0								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Female	30	34	Hoth 1.5mm	Other medium	24 hr			2	1	50								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Male	35	39	Hoth 1.5mm	Other medium	24 hr			2	1	50								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Female	35	39	Hoth 1.5mm	Other medium	24 hr			0	0	0								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Male	40	44	Hoth 1.5mm	Other medium	24 hr			3	0	0								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Female	40	44	Hoth 1.5mm	Other medium	24 hr			1	1	100								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Male	45	49	Hoth 1.5mm	Other medium	24 hr			0	0	0								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Female	45	49	Hoth 1.5mm	Other medium	24 hr			0	0	0								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Male	50	+	Hoth 1.5mm	Other medium	24 hr			1	1	100								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Female	50	+	Hoth 1.5mm	Other medium	24 hr			0	0	0								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Male	All	All	Hoth 1.5mm	Other medium	24 hr			21	13	61.904762								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Female	All	All	Hoth 1.5mm	Other medium	24 hr			7	6	85.714286								
SOH258	325	2009.06.30	SOH	1980	1990	Forest	Average	Group of villages	-Saharan Africa, W	Liberia	Margibi	Firestone District	Division 36	6.267011	-10.342426		S. yahense	S. soubrense				Both	All	All	Hoth 1.5mm	Other medium	24 hr			56	35	62.5								
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	0	4	Hoth 1.5mm	Other medium	24 hr			2	0	0				Arithmetic	mf/ss	Positives	0	0
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	0	4	Hoth 1.5mm	Other medium	24 hr			2	0	0				Arithmetic	mf/ss	Positives	0	0
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	5	9	Hoth 1.5mm	Other medium	24 hr			53	30	56.603774				Arithmetic	mf/ss	Positives	30	13.9
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	5	9	Hoth 1.5mm	Other medium	24 hr			43	29	67.44186				Arithmetic	mf/ss	Positives	29	6
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	10	14	Hoth 1.5mm	Other medium	24 hr			76	59	77.631579				Arithmetic	mf/ss	Positives	59	12.7
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	10	14	Hoth 1.5mm	Other medium	24 hr			43	35	81.395349				Arithmetic	mf/ss	Positives	35	6.5
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	15	19	Hoth 1.5mm	Other medium	24 hr			38	32	84.210526				Arithmetic	mf/ss	Positives	32	18.7
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	15	19	Hoth 1.5mm	Other medium	24 hr			20	16	80				Arithmetic	mf/ss	Positives	16	8.2
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	20	24	Hoth 1.5mm	Other medium	24 hr			35	30	85.714286				Arithmetic	mf/ss	Positives	30	26.4
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	20	24	Hoth 1.5mm	Other medium	24 hr			16	12	75				Arithmetic	mf/ss	Positives	12	10.8
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	25	29	Hoth 1.5mm	Other medium	24 hr			36	32	88.888889				Arithmetic	mf/ss	Positives	32	22.6
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	25	29	Hoth 1.5mm	Other medium	24 hr			26	22	84.615385				Arithmetic	mf/ss	Positives	22	19.5
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	30	34	Hoth 1.5mm	Other medium	24 hr			44	42	95.454545				Arithmetic	mf/ss	Positives	42	31.8
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	30	34	Hoth 1.5mm	Other medium	24 hr			26	21	80.769231				Arithmetic	mf/ss	Positives	21	21.3
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	35	39	Hoth 1.5mm	Other medium	24 hr			32	28	87.5				Arithmetic	mf/ss	Positives	28	21.4
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	35	39	Hoth 1.5mm	Other medium	24 hr			10	9	90				Arithmetic	mf/ss	Positives	9	18.9
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	40	44	Hoth 1.5mm	Other medium	24 hr			24	19	79.166667				Arithmetic	mf/ss	Positives	19	16
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	40	44	Hoth 1.5mm	Other medium	24 hr			10	10	100				Arithmetic	mf/ss	Positives	10	19.4
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	45	49	Hoth 1.5mm	Other medium	24 hr			17	15	88.235294				Arithmetic	mf/ss	Positives	15	29.6
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	45	49	Hoth 1.5mm	Other medium	24 hr			4	4	100				Arithmetic	mf/ss	Positives	4	25.5
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	50	54	Hoth 1.5mm	Other medium	24 hr			17	17	100				Arithmetic	mf/ss	Positives	17	30.8
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	50	54	Hoth 1.5mm	Other medium	24 hr			9	9	100				Arithmetic	mf/ss	Positives	9	20.6
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	All	All	Hoth 1.5mm	Other medium	24 hr			374	304	81.283422				Arithmetic	mf/ss	Positives	304	22.2
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	All	All	Hoth 1.5mm	Other medium	24 hr			209	167	79.904306				Arithmetic	mf/ss	Positives	167	28.4
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	0	4	Hoth 1.5mm	Other medium	24 hr			2	0	0				Arithmetic	mf/mg	Positives	0	0
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	0	4	Hoth 1.5mm	Other medium	24 hr			2	0	0				Arithmetic	mf/mg	Positives	0	0
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	5	9	Hoth 1.5mm	Other medium	24 hr			53	30	56.603774				Arithmetic	mf/mg	Positives	30	16.6
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	5	9	Hoth 1.5mm	Other medium	24 hr			43	29	67.44186				Arithmetic	mf/mg	Positives	29	7.2
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	10	14	Hoth 1.5mm	Other medium	24 hr			76	59	77.631579				Arithmetic	mf/mg	Positives	59	15.2
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	10	14	Hoth 1.5mm	Other medium	24 hr			43	35	81.395349				Arithmetic	mf/mg	Positives	35	7.8
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	15	19	Hoth 1.5mm	Other medium	24 hr			38	32	84.210526				Arithmetic	mf/mg	Positives	32	22.4
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	15	19	Hoth 1.5mm	Other medium	24 hr			20	16	80				Arithmetic	mf/mg	Positives	16	9.8
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	20	24	Hoth 1.5mm	Other medium	24 hr			35	30	85.714286				Arithmetic	mf/mg	Positives	30	31.6
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	20	24	Hoth 1.5mm	Other medium	24 hr			16	12	75				Arithmetic	mf/mg	Positives	12	12.9
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	25	29	Hoth 1.5mm	Other medium	24 hr			36	32	88.888889				Arithmetic	mf/mg	Positives	32	27.1
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	25	29	Hoth 1.5mm	Other medium	24 hr			26	22	84.615385				Arithmetic	mf/mg	Positives	22	23.4
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	30	34	Hoth 1.5mm	Other medium	24 hr			44	42	95.454545				Arithmetic	mf/mg	Positives	42	58.1
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	30	34	Hoth 1.5mm	Other medium	24 hr			26	21	80.769231				Arithmetic	mf/mg	Positives	21	25.5
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	35	39	Hoth 1.5mm	Other medium	24 hr			32	28	87.5				Arithmetic	mf/mg	Positives	28	25.6
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	35	39	Hoth 1.5mm	Other medium	24 hr			10	9	90				Arithmetic	mf/mg	Positives	9	22.6
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	40	44	Hoth 1.5mm	Other medium	24 hr			24	19	79.166667				Arithmetic	mf/mg	Positives	19	43.1
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	40	44	Hoth 1.5mm	Other medium	24 hr			10	10	100				Arithmetic	mf/mg	Positives	10	23.2
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	45	49	Hoth 1.5mm	Other medium	24 hr			17	15	88.235294				Arithmetic	mf/mg	Positives	15	35.4
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	45	49	Hoth 1.5mm	Other medium	24 hr			4	4	100				Arithmetic	mf/mg	Positives	4	30.5
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	50	54	Hoth 1.5mm	Other medium	24 hr			17	17	100				Arithmetic	mf/mg	Positives	17	36.9
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	50	54	Hoth 1.5mm	Other medium	24 hr			9	9	100				Arithmetic	mf/mg	Positives	9	24.7
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Male	All	All	Hoth 1.5mm	Other medium	24 hr			374	304	81.283422				Arithmetic	mf/mg	Positives	304	28.4
SOH259	325	2009.06.30	SOH	1980	1990	Forest	Average	District	-Saharan Africa, W	Liberia	Margibi	Firestone District	All Divisions	6.267011	-10.342426		S. yahense	S. soubrense				Female	All	All	Hoth 1.5mm	Other medium	24 hr			209	167	79.904306				Arithmetic	mf/mg	Positives	167	17.1
SOH260	313	2009.07.03	SOH	1982	1990	Savannah	Dry	Region	-Saharan Africa, E	Sudan	Abu Hamed			19.52873	33.288574		S. dammousum s.l.	S. sirbanum				Both	10	90	Walzer punch	Water	1-6 hr			208	60	28.846154	208	26	12.5	Arithmetic	mf/mg	Positives	60	2.9
SOH261	310	2009.07.31	SOH	1985	1990	Savannah	Dry	Village	-Saharan Africa, W	Nigeria	Kwara	Shagunu	Amboshidi	10.352072	4.456329		S. dammousum s.l.					Male	All	All	Stolz punch	Saline	6 - 12 hr			54	22.95	42.5								
SOH261	310	2009.07.31	SOH	1985	1990	Savannah	Dry	Village	-Saharan Africa, W	Nigeria	Kwara	Shagunu	Amboshidi	10.352072	4.456329		S. dammousum s.l.					Female	All	All	Stolz punch	Saline	6 - 12 hr			42	8.988	21.4								
SOH261	310	2009.07.31	SOH	1985	1990	Savannah	Dry	Village	-Saharan Africa, W	Nigeria	Kwara	Shagunu	Amboshidi	10.352072	4.456329		S. dammousum s.l.					Both	All	All	Stolz punch	Saline	6 - 12 hr			96	31.968	33.3								
SOH262	310	2009.07.31	SOH	1985	1990	Savannah	Dry	Village	-Saharan Africa, W	Nigeria	Kwara	Shagunu	Luma I	10.352072	4.456329		S. dammousum s.l.					Male	All	All	Stolz punch	Saline	6 - 12 hr			41	28.003	68.3								
SOH262	310	2009.07.31	SOH	1985	1990	Savannah	Dry	Village	-Saharan Africa, W	Nigeria	Kwara	Shagunu	Luma I	10.352072	4.456329		S. dammousum s.l.					Female	All	All	Stolz punch	Saline	6 - 12 hr			42	17.01	40.5								
SOH262	310	2009.07.31	SOH	1985	1990	Savannah	Dry	Village	-Saharan Africa, W	Nigeria	Kwara	Shagunu	Luma I	10.352072	4.456329		S. dammousum s.l.					Both	All	All	Stolz punch	Saline	6 - 12 hr			83	44.986	54.2								
SOH263	310	2009.07.31	SOH	1985	1990	Savannah	Dry	Village	-Saharan Africa, W	Nigeria	Kwara	Shagunu	Luma II	10.352072	4.456329		S. dammousum s.l.					Male	All	All	Stolz punch	Saline	6 - 12 hr			96	40.992	42.7								
SOH263	310	2009.07.31	SOH	1985	1990	Savannah	Dry	Village	-Saharan Africa, W	Nigeria	Kwara	Shagunu	Luma II	10.352072	4.456329		S. dammousum s.l.					Female	All	All	Stolz punch	Saline	6 - 12 hr			87	20.967	24.1								
SOH263	310	2009.07.31	SOH	1985	1990	Savannah	Dry	Village	-Saharan Africa, W	Nigeria	Kwara	Shagunu	Luma II	10.352072	4.456329		S. dammousum s.l.					Both	All	All	Stolz punch	Saline	6 - 12 hr			183	62.037	33.9								
SOH264	310	2009.07.31

Study Site ID	Publication ID	Date entered	Curator	Study year	GBD estimate year	Bioclime	Season	Settlement Type	GBD Region	Country	Region / State	District / Ward / Municipality/River Basin	Location name	Latitude	Longitude	Elevation (m)	Primary vector	Secondary vector	Biting rate	Biting rate 95% CI	Population size	Sex	Lower age	Upper age	Diagnostic method	Incubation media	Incubation time	Endemicity level	Treatment rounds	Coverage est.	ss_exam	ss_pos	ss_prev	nod_exam	nod_pos	nod_prev	mf_measure	mf_unit	mf_denom	mf_exam	mf_prev
SOH275	310	2009.08.03	SOH	1985	1990	Savannah	Dry	Group of villages	-Saharan Africa, Wa	Nigeria	Kwara	Shagunu	All Villages	10.352072	4.456329		S. damnosum s.l.					Male	40	49	Stolz punch	Saline	6 - 12 hr				148	60.976	41.2								
SOH275	310	2009.08.03	SOH	1985	1990	Savannah	Dry	Group of villages	-Saharan Africa, Wa	Nigeria	Kwara	Shagunu	All Villages	10.352072	4.456329		S. damnosum s.l.					Female	40	49	Stolz punch	Saline	6 - 12 hr				110	36.96	33.6								
SOH275	310	2009.08.03	SOH	1985	1990	Savannah	Dry	Group of villages	-Saharan Africa, Wa	Nigeria	Kwara	Shagunu	All Villages	10.352072	4.456329		S. damnosum s.l.					Male	50	59	Stolz punch	Saline	6 - 12 hr				120	56.04	46.7								
SOH275	310	2009.08.03	SOH	1985	1990	Savannah	Dry	Group of villages	-Saharan Africa, Wa	Nigeria	Kwara	Shagunu	All Villages	10.352072	4.456329		S. damnosum s.l.					Female	50	59	Stolz punch	Saline	6 - 12 hr				96	30.048	31.3								
SOH275	310	2009.08.03	SOH	1985	1990	Savannah	Dry	Group of villages	-Saharan Africa, Wa	Nigeria	Kwara	Shagunu	All Villages	10.352072	4.456329		S. damnosum s.l.					Male	60	69	Stolz punch	Saline	6 - 12 hr				94	47	50								
SOH275	310	2009.08.03	SOH	1985	1990	Savannah	Dry	Group of villages	-Saharan Africa, Wa	Nigeria	Kwara	Shagunu	All Villages	10.352072	4.456329		S. damnosum s.l.					Female	60	69	Stolz punch	Saline	6 - 12 hr				78	32.994	42.3								
SOH275	310	2009.08.03	SOH	1985	1990	Savannah	Dry	Group of villages	-Saharan Africa, Wa	Nigeria	Kwara	Shagunu	All Villages	10.352072	4.456329		S. damnosum s.l.					Male	70	+	Stolz punch	Saline	6 - 12 hr				73	31.025	42.5								
SOH275	310	2009.08.03	SOH	1985	1990	Savannah	Dry	Group of villages	-Saharan Africa, Wa	Nigeria	Kwara	Shagunu	All Villages	10.352072	4.456329		S. damnosum s.l.					Female	70	+	Stolz punch	Saline	6 - 12 hr				53	18.974	35.8								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Male	0	9	Holth 1.5mm	Saline	6 - 12 hr				159	9	5.6602774								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Female	0	9	Holth 1.5mm	Saline	6 - 12 hr				161	0	0								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Both	0	9	Holth 1.5mm	Saline	6 - 12 hr				320	9	2.8125								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Male	10	19	Holth 1.5mm	Saline	6 - 12 hr				61	8	13.114754								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Female	10	19	Holth 1.5mm	Saline	6 - 12 hr				119	4	3.3613445								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Both	10	19	Holth 1.5mm	Saline	6 - 12 hr				180	12	6.6666667								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Male	20	29	Holth 1.5mm	Saline	6 - 12 hr				92	27	29.347826								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Female	20	29	Holth 1.5mm	Saline	6 - 12 hr				164	20	12.195122								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Both	20	29	Holth 1.5mm	Saline	6 - 12 hr				256	47	18.359375								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Male	30	39	Holth 1.5mm	Saline	6 - 12 hr				173	63	36.416185								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Female	30	39	Holth 1.5mm	Saline	6 - 12 hr				158	31	19.620253								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Both	30	39	Holth 1.5mm	Saline	6 - 12 hr				331	94	28.398792								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Male	40	49	Holth 1.5mm	Saline	6 - 12 hr				103	41	39.805825								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Female	40	49	Holth 1.5mm	Saline	6 - 12 hr				65	13	20								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Both	40	49	Holth 1.5mm	Saline	6 - 12 hr				168	54	32.142857								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Male	50	59	Holth 1.5mm	Saline	6 - 12 hr				39	31	79.487179								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Female	50	59	Holth 1.5mm	Saline	6 - 12 hr				36	4	11.111111								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Both	50	59	Holth 1.5mm	Saline	6 - 12 hr				75	35	46.666667								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Male	60	69	Holth 1.5mm	Saline	6 - 12 hr				22	14	63.636364								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Female	60	69	Holth 1.5mm	Saline	6 - 12 hr				18	2	11.111111								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Both	60	69	Holth 1.5mm	Saline	6 - 12 hr				40	16	40								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Male	All	All	Holth 1.5mm	Saline	6 - 12 hr				647	193	29.829985								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Female	All	All	Holth 1.5mm	Saline	6 - 12 hr				721	74	10.263523								
SOH276	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Group of villages	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	All Villages	14.280388	36.61749	500-1000	S. damnosum s.l.				14242	Both	All	All	Holth 1.5mm	Saline	6 - 12 hr				1370	267	19.489051	1370	21	1.53284672	Arithmetic	mf/mg	Positives	1370	3.6
SOH277	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Village	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	Bewal	14.280388	36.61749	500-1000	S. damnosum s.l.				1002	Both	All	All	Holth 1.5mm	Saline	6 - 12 hr				93	43	46.236559				Arithmetic	mf/mg	Positives	93	3.6
SOH278	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Village	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	Humera	14.280388	36.61749	500-1000	S. damnosum s.l.				10000	Both	All	All	Holth 1.5mm	Saline	6 - 12 hr				967	133	13.753878				Arithmetic	mf/mg	Positives	967	4.1
SOH279	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Village	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	Baaker	14.280388	36.61749	500-1000	S. damnosum s.l.				1170	Both	All	All	Holth 1.5mm	Saline	6 - 12 hr				115	43	37.391304				Arithmetic	mf/mg	Positives	115	4.1
SOH280	299	2009.08.03	SOH	1984	1990	Savannah	Unknown	Village	- Saharan Africa, E	Ethiopia	Mi'eraba wei	Tigray	Helgum + Umhajer	14.280388	36.61749	500-1000	S. damnosum s.l.				2070	Both	All	All	Holth 1.5mm	Saline	6 - 12 hr				191	48	25.13089				Arithmetic	mf/mg	Positives	191	4.6
SOH281	297	2009.08.03	SOH	1986	1990	Savannah	Unknown	District	-Saharan Africa, Wa	Nigeria	Kwara	Agwara	11 Villages	10.352072	4.456329		S. damnosum s.l.					Both	All	All	Stolz punch	Saline	6 - 12 hr				1060	103	9.716981								
SOH282	297	2009.08.03	SOH	1986	1990	Savannah	Unknown	District	-Saharan Africa, Wa	Nigeria	Kwara	Babana	12 Villages	10.352072	4.456329		S. damnosum s.l.					Both	All	All	Stolz punch	Saline	6 - 12 hr				993	483	48.640483								
SOH283	297	2009.08.03	SOH	1986	1990	Savannah	Unknown	District	-Saharan Africa, Wa	Nigeria	Kwara	Ile-Ire	15 Villages	10.352072	4.456329		S. damnosum s.l.					Both	All	All	Stolz punch	Saline	6 - 12 hr				2468	1540	62.398703								
SOH284	297	2009.08.03	SOH	1986	1990	Savannah	Unknown	District	-Saharan Africa, Wa	Nigeria	Kwara	Shagunu	13 Villages	10.352072	4.456329		S. damnosum s.l.					Both	All	All	Stolz punch	Saline	6 - 12 hr				1887	538	28.510864								
SOH285	297	2009.08.03	SOH	1986	1990	Savannah	Unknown	District	-Saharan Africa, Wa	Nigeria	Kwara	Moro	Shao	10.352072	4.456329		S. damnosum s.l.					Both	All	All	Stolz punch	Saline	6 - 12 hr				2018	1101	54.558969								
SOH286	297	2009.08.03	SOH	1986	1990	Savannah	Unknown	Region	-Saharan Africa, Wa	Nigeria	Kwara	All Districts	All Villages	10.352072	4.456329		S. damnosum s.l.					Male	All	All	Stolz punch	Saline	6 - 12 hr				4125	2088	50.618182								
SOH286	297	2009.08.03	SOH	1986	1990	Savannah	Unknown	Region	-Saharan Africa, Wa	Nigeria	Kwara	All Districts	All Villages	10.352072	4.456329		S. damnosum s.l.					Female	All	All	Stolz punch	Saline	6 - 12 hr				4301	1677	38.990932								
SOH286	297	2009.08.03	SOH	1986	1990	Savannah	Unknown	Region	-Saharan Africa, Wa	Nigeria	Kwara	All Districts	All Villages	10.352072	4.456329		S. damnosum s.l.					Both	0	<15	Stolz punch	Saline	6 - 12 hr				2106	143	6.7901235								
SOH286	297	2009.08.03	SOH	1986	1990	Savannah	Unknown	Region	-Saharan Africa, Wa	Nigeria	Kwara	All Districts	All Villages	10.352072	4.456329		S. damnosum s.l.					Both	15	+	Stolz punch	Saline	6 - 12 hr				6320	3622	57.310127								
SOH286	297	2009.08.03	SOH	1986	1990	Savannah	Unknown	Region	-Saharan Africa, Wa	Nigeria	Kwara	All Districts	All Villages	10.352072	4.456329		S. damnosum s.l.					Both	All	All	Stolz punch	Saline	6 - 12 hr				8426	3765	44.681124								
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Village	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	M. Sundus	13.25	36.13333						550	Both	All	All	Skin snip	Other medium	1-6 hr				106	44.626	42.1	121	37	30.5785124					
SOH288	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Village	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	M. Hagar el Zurrug	13.216667	36.15						200	Both	All	All							51	32.334	63.4	52	14	26.9230769					
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	1	6	Skin snip	Other medium	1-6 hr							16	0	0					
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	7	12	Skin snip	Other medium	1-6 hr							36	5	13.8888889					
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	13	19	Skin snip	Other medium	1-6 hr							39	10	25.6410256					
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	20	29	Skin snip	Other medium	1-6 hr							23	6	26.0869565					
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	30	39	Skin snip	Other medium	1-6 hr							22	8	36.3636364					
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	40	49	Skin snip	Other medium	1-6 hr							14	10	71.4285714					
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	50	+	Skin snip	Other medium	1-6 hr							23	12	52.173913					
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	All	All	Skin snip	Other medium	1-6 hr				157			173	51	29.4797688					
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	All	All	Skin snip	Other medium	1-6 hr										Arithmetic	mf/mg	Positives	34	0.1-6
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	All	All	Skin snip	Other medium	1-6 hr										Arithmetic	mf/mg	Positives	32	7-10
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	All	All	Skin snip	Other medium	1-6 hr										Arithmetic	mf/mg	Positives	8	11-20
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	All	All	Skin snip	Other medium	1-6 hr										Arithmetic	mf/mg	Positives	3	21-30
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	All	All	Skin snip	Other medium	1-6 hr										Arithmetic	mf/mg	Positives	1	31-40
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	All	All	Skin snip	Other medium	1-6 hr										Arithmetic	mf/mg	Positives	3	41-60
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	All	All	Skin snip	Other medium	1-6 hr										Arithmetic	mf/mg	Positives	2	61-80
SOH287	294	2009.08.07	SOH	1985	1990	Savannah	Dry	Group of villages	- Saharan Africa, E	Sudan	Kasalla Province	Gedaref Health district	All Villages	13.2325	36.1415						750	Both	All	All	Skin snip	Other medium	1-6 hr										Arithmetic	mf/mg	Positives	1	>81

Appendix F. Papers Published During PhD Registration

F.1. Paper from Chapter 3

PLOS | NEGLECTED TROPICAL DISEASES

RESEARCH ARTICLE

Model-Based Geostatistical Mapping of the Prevalence of *Onchocerca volvulus* in West Africa

Simon J. O'Hanlon[1]*, Hannah C. Slater[1,2], Robert A. Cheke[1,3], Boakye A. Boatin[4], Luc E. Coffeng[5], Sébastien D. S. Pion[6], Michel Boussinesq[6], Honorat G. M. Zouré[7], Wilma A. Stolk[5], María-Gloria Basáñez[1,8]*

1 Department of Infectious Disease Epidemiology, School of Public Health, Faculty of Medicine (St Mary's Campus), Imperial College London, London, United Kingdom, **2** MRC Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom, **3** Natural Resources Institute, University of Greenwich at Medway, Chatham, Kent, United Kingdom, **4** Lymphatic Filariasis Support Centre, Department of Parasitology, Noguchi Memorial Institute for Medical Research, University of Ghana, Legon, Ghana, **5** Department of Public Health, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands, **6** UMI 233, Institut de Recherche pour le Développement (IRD) and University of Montpellier 1, Montpellier, France, **7** African Programme for Onchocerciasis Control (APOC), World Health Organization (WHO), Ouagadougou, Burkina Faso, **8** London Centre for Neglected Tropical Disease Research, Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom

* simon.ohanlon@gmail.com (SJO); m.basanez@imperial.ac.uk (MGB)

OPEN ACCESS

Citation: O'Hanlon SJ, Slater HC, Cheke RA, Boatin BA, Coffeng LE, Pion SDS, et al. (2016) Model-Based Geostatistical Mapping of the Prevalence of *Onchocerca volvulus* in West Africa. PLoS Negl Trop Dis 10(1): e0004328. doi:10.1371/journal.pntd.0004328

Editor: Ricardo J. Soares Magalhaes, University of Queensland School of Veterinary Science, AUSTRALIA

Received: November 21, 2014

Accepted: December 4, 2015

Published: January 15, 2016

Copyright: © 2016 O'Hanlon et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data are freely available, excepting the original epidemiological data, belonging to and collected under the auspices of the Onchocerciasis Control Programme (OCP). Since the cessation of the OCP, both the World Health Organization (WHO) and the African Programme for Onchocerciasis Control (APOC) are guardians of these data. Any request for these data should therefore be addressed to the current Director of APOC, Dr Jean-Baptiste Roungou at roungouj@who.int.

Abstract

Background

The initial endemicity (pre-control prevalence) of onchocerciasis has been shown to be an important determinant of the feasibility of elimination by mass ivermectin distribution. We present the first geostatistical map of microfilarial prevalence in the former Onchocerciasis Control Programme in West Africa (OCP) before commencement of antivectorial and anti-parasitic interventions.

Methods and Findings

Pre-control microfilarial prevalence data from 737 villages across the 11 constituent countries in the OCP epidemiological database were used as ground-truth data. These 737 data points, plus a set of statistically selected environmental covariates, were used in a Bayesian model-based geostatistical (B-MBG) approach to generate a continuous surface (at pixel resolution of 5 km x 5km) of microfilarial prevalence in West Africa prior to the commencement of the OCP. Uncertainty in model predictions was measured using a suite of validation statistics, performed on bootstrap samples of held-out validation data. The mean Pearson's correlation between observed and estimated prevalence at validation locations was 0.693; the mean prediction error (average difference between observed and estimated values) was 0.77%, and the mean absolute prediction error (average magnitude of difference between observed and estimated values) was 12.2%. Within OCP boundaries, 17.8 million

F.2. Papers not directly related to PhD

OPEN ACCESS Freely available online

PLOS | NEGLECTED TROPICAL DISEASES

Viewpoints

The Global Burden of Disease Study 2010: Interpretation and Implications for the Neglected Tropical Diseases

Peter J. Hotez[1,2,3]*, Miriam Alvarado[4], María-Gloria Basáñez[5], Ian Bolliger[4], Rupert Bourne[6], Michel Boussinesq[7], Simon J. Brooker[8], Ami Shah Brown[9], Geoffrey Buckle[10], Christine M. Budke[11], Hélène Carabin[12], Luc E. Coffeng[4,13], Eric M. Fèvre[14,15], Thomas Fürst[5,16,17], Yara A. Halasa[18], Rashmi Jasrasaria[4], Nicole E. Johns[4], Jennifer Keiser[16,17], Charles H. King[19], Rafael Lozano[4], Michele E. Murdoch[20], Simon O'Hanlon[5], Sébastien D. S. Pion[7], Rachel L. Pullan[8], Kapa D. Ramaiah[21], Thomas Roberts[22], Donald S. Shepard[18], Jennifer L. Smith[8], Wilma A. Stolk[13], Eduardo A. Undurraga[18], Jürg Utzinger[16,17], Mengru Wang[4], Christopher J. L. Murray[4]*, Mohsen Naghavi[4]*

1 National School of Tropical Medicine at Baylor College of Medicine, Houston, Texas, United States of America, **2** Sabin Vaccine Institute and Texas Children's Hospital Center for Vaccine Development, Houston, Texas, United States of America, **3** James A. Baker III Institute at Rice University, Houston, Texas, United States of America, **4** Institute for Health Metrics and Evaluation, University of Washington, Seattle, Washington, United States of America, **5** Imperial College London, London, United Kingdom, **6** Vision and Eye Research Unit, Anglia Ruskin University, Cambridge, United Kingdom, **7** Institut de Recherche pour le Développement, Montpellier, France, **8** London School of Hygiene and Tropical Medicine, London, United Kingdom, **9** Inovio Pharmaceuticals, Inc., Blue Bell, Pennsylvania, United States of America, **10** Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, United States of America, **11** Texas A&M University, College Station, Texas, United States of America, **12** University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, United States of America, **13** Erasmus MC, University Medical Center Rotterdam, Rotterdam, Netherlands, **14** Institute of Infection and Global Health, University of Liverpool, Liverpool, United Kingdom, **15** International Livestock Research Institute, Nairobi, Kenya, **16** Swiss Tropical and Public Health Institute, Basel, Switzerland, **17** University of Basel, Basel, Switzerland, **18** Brandeis University, Waltham, Massachusetts, United States of America, **19** Case Western Reserve University, Cleveland, Ohio, United States of America, **20** Watford General Hospital, Watford, United Kingdom, **21** Vector Control Research Centre, Pondicherry, India, **22** Stanford University School of Medicine, Stanford, California, United States of America

Introduction

The publication of the Global Burden of Disease Study 2010 (GBD 2010) and the accompanying collection of *Lancet* articles in December 2012 provided the most comprehensive attempt to quantify the burden of almost 300 diseases, injuries, and risk factors, including neglected tropical diseases (NTDs) [1–3]. The disability-adjusted life year (DALY), the metric used in the GBD 2010, is a tool which may be used to assess and compare the relative impact of a number of diseases locally and globally [4–6]. Table 1 lists the major NTDs as defined by the World Health Organization (WHO) [7] and their estimated DALYs [1]. With a few exceptions, most of the NTDs currently listed by the WHO [7] or those on the expanded list from *PLOS Neglected Tropical Diseases* [8] are disablers rather than killers, so the DALY estimates represent one of the few metrics available that could fully embrace the chronic effects of these infections.

Even DALYs, however, do not tell the complete story of the harmful effects from NTDs. Some of the specific and potential shortcomings of GBD 2010 have been highlighted elsewhere [9]. Furthermore, DALYs measure only direct health loss and, for example, do not consider the economic impact of the NTDs that results from detrimental effects on school attendance and child development, agriculture (especially from zoonotic NTDs), and overall economic productivity [10,11]. Nor do DALYs account for direct costs of treatment, surveillance, and prevention measures. Yet, economic impact has emerged as an essential feature of the NTDs, which may trap people in a cycle of poverty and disease [10–12]. Additional aspects not considered by the DALY metrics are the important elements of social stigma for many of the NTDs and the spillover effects to family and community members [13,14], loss of tourism [15], and health system overload (e.g., during dengue outbreaks). Ultimately NTD control and elimination efforts could produce social and economic benefits not necessarily reflected in the DALY metrics, especially among the most affected poor communities [11].

Variations in DALYs

Despite the importance of the concept of disease burden and disability to the NTD community, assigning DALYs or related metrics to each NTD has been a bit of a roller-coaster ride over the past decade and may continue to be for many years to come. Significant variations in ascribing DALYs to the NTDs are due to many factors, including data scarcity and inherent difficulties in accurately estimating the number of individuals at risk, the number of incident cases, the number of prevalent cases, and, among these, the duration of the infection. Challenges also include uncertainty about the relationship between acute and chronic infections and their link to specific morbidities, duration of morbidity, and the proportion of the population infected or with morbidities

Citation: Hotez PJ, Alvarado M, Basáñez M-G, Bolliger I, Bourne R, et al. (2014) The Global Burden of Disease Study 2010: Interpretation and Implications for the Neglected Tropical Diseases. PLoS Negl Trop Dis 8(7): e2865. doi:10.1371/journal.pntd.0002865

Editor: Nilanthi de Silva, University of Kelaniya, Sri Lanka

Published July 24, 2014

Copyright: © 2014 Hotez et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors have indicated that no funding was received for this work.

Competing Interests: I have read the journal's policy and have the following conflicts: ASB is employed by a commercial company, Inovio Pharmaceuticals. This does not alter our adherence to all PLOS NTDs policies on sharing data and materials. There are no other competing interests to declare.

* Email: hotez@bcm.edu (PJH); cjlm@uw.edu (CJLM); nagham@uw.edu (MN)

These authors contributed equally to this work.

OPEN ACCESS Freely available online

PLOS | NEGLECTED TROPICAL DISEASES

Onchocerciasis: The Pre-control Association between Prevalence of Palpable Nodules and Skin Microfilariae

Luc E. Coffeng[1*], Sébastien D. S. Pion[2], Simon O'Hanlon[3], Simon Cousens[4], Adenike O. Abiose[5], Peter U. Fischer[6,7], Jan H. F. Remme[8], K. Yankum Dadzie[9], Michele E. Murdoch[10], Sake J. de Vlas[1], María-Gloria Basáñez[3], Wilma A. Stolk[1], Michel Boussinesq[2]

1 Department of Public Health, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands, **2** UMI 233, Institut de Recherche pour le Développement (IRD), University of Montpellier 1, Montpellier, France, **3** Department of Infectious Disease Epidemiology, School of Public Health, Faculty of Medicine (St Mary's Campus), Imperial College London, London, United Kingdom, **4** Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom, **5** Sightcare International, Secretariat Main Office, Ibadan, Oyo State, Nigeria, **6** Washington University School of Medicine, Infectious Disease Division, St. Louis, Missouri, United States of America, **7** Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany, **8** Independent Consultant, Ornex, France, **9** Independent Consultant, Accra, Ghana, **10** Department of Dermatology, Watford General Hospital, Watford, Hertfordshire, United Kingdom

Abstract

Background: The prospect of eliminating onchocerciasis from Africa by mass treatment with ivermectin has been rejuvenated following recent successes in foci in Mali, Nigeria and Senegal. Elimination prospects depend strongly on local transmission conditions and therefore on pre-control infection levels. Pre-control infection levels in Africa have been mapped largely by means of nodule palpation of adult males, a relatively crude method for detecting infection. We investigated how informative pre-control nodule prevalence data are for estimating the pre-control prevalence of microfilariae (mf) in the skin and discuss implications for assessing elimination prospects.

Methods and Findings: We analyzed published data on pre-control nodule prevalence in males aged ≥20 years and mf prevalence in the population aged ≥5 years from 148 African villages. A meta-analysis was performed by means of Bayesian hierarchical multivariate logistic regression, accounting for measurement error in mf and nodule prevalence, bioclimatic zones, and other geographical variation. There was a strong positive correlation between nodule prevalence in adult males and mf prevalence in the general population. In the forest-savanna mosaic area, the pattern in nodule and mf prevalence differed significantly from that in the savanna or forest areas.

Significance: We provide a tool to convert pre-control nodule prevalence in adult males to mf prevalence in the general population, allowing historical data to be interpreted in terms of elimination prospects and disease burden of onchocerciasis. Furthermore, we identified significant geographical variation in mf prevalence and nodule prevalence patterns warranting further investigation of geographical differences in transmission patterns of onchocerciasis.

Citation: Coffeng LE, Pion SDS, O'Hanlon S, Cousens S, Abiose AO, et al. (2013) Onchocerciasis: The Pre-control Association between Prevalence of Palpable Nodules and Skin Microfilariae. PLoS Negl Trop Dis 7(4): e2168. doi:10.1371/journal.pntd.0002168

Editor: Christian Bottomley, London School of Hygiene and Tropical Medicine, United Kingdom

Received June 8, 2012; **Accepted** March 1, 2013; **Published** April 11, 2013

Copyright: © 2013 Coffeng et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: LEC and WAS were funded by a research grant from the African Programme for Onchocerciasis Control (APOC, www.who.int/apoc) for a Health Impact Assessment of APOC (CEV/APOC/241/08/TSA). MGB and WAS acknowledge seed funding from the Institute of Health Metrics and Evaluation (Seattle, USA, www.healthmetricsandevaluation.org/) as co-disease leaders for the project on global burden of onchocercal disease. SOH was funded by an ESRC (Economics and Social Sciences Research Council, UK, www.esrc.ac.uk) PhD studentship. Surveys in Cameroon were performed with the financial supports from the Special Programme for Research and Training in Tropical Diseases (Vina; www.who.int/tdr), Helen Keller International (Lekie; www.hki.org) and Institut de Recherche pour le Développement (Mbam and Faro; www.ird.fr). Surveys in Nigeria were supported by WHO/UNDP/World Bank Special Programme for Research and Training on Tropical Diseases (Project ID Numbers 870456 and 910553), the Leverhulme Trust and Sightsavers. MGB acknowledges funding from the Wellcome Trust (http://www.wellcome.ac.uk, project grant 092677/Z/10/Z). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: l.coffeng@erasmusmc.nl

These authors contributed equally to this work.

These authors also contributed equally to this work.

Introduction

In 1995, the World Health Organization launched the African Programme for Onchocerciasis Control (APOC). At that time, APOC aimed to control morbidity due to onchocerciasis (river blindness) in Africa, with a focus on those countries not covered by the previous Onchocerciasis Control Programme in West Africa (OCP). Since 1995, APOC has successfully coordinated mass treatment with ivermectin in sixteen onchocerciasis-endemic African countries [1]. Until recently, elimination of onchocerciasis from African foci was deemed to be not achievable by means of mass ivermectin treatment alone, considering the large size of the transmission zones, the mobility of the insect vectors and human

Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010

Christopher J L Murray†‡, Theo Vos, Rafael Lozano, Mohsen Naghavi, Abraham D Flaxman, Catherine Michaud, Majid Ezzati, Kenji Shibuya, Joshua A Salomon, Safa Abdalla*, Victor Aboyans*, Jerry Abraham*, Ilana Ackerman*, Rakesh Aggarwal*, Stephanie Y Ahn*, Mohammed K Ali*, Mohammad A AlMazroa*, Miriam Alvarado*, H Ross Anderson*, Laurie M Anderson*, Kathryn G Andrews*, Charles Atkinson*, Larry M Baddour*, Adil N Bahalim*, Suzanne Barker-Collo*, Lope H Barrero*, David H Bartels*, Maria-Gloria Basáñez*, Amanda Baxter*, Michelle L Bell*, Emelia J Benjamin*, Derrick Bennett*, Eduardo Bernabé*, Kavi Bhalla*, Bishal Bhandari*, Boris Bikbov*, Aref Bin Abdulhak*, Gretchen Birbeck*, James A Black*, Hannah Blencowe*, Jed D Blore*, Fiona Blyth*, Ian Bolliger*, Audrey Bonaventure*, Soufiane Boufous*, Rupert Bourne*, Michel Boussinesq*, Tasanee Braithwaite*, Carol Brayne*, Lisa Bridgett*, Simon Brooker*, Peter Brooks*, Traolach S Brugha*, Claire Bryan-Hancock*, Chiara Bucello*, Rachelle Buchbinder*, Geoffrey Buckle*, Christine M Budke*, Michael Burch*, Peter Burney*, Roy Burstein*, Bianca Calabria*, Benjamin Campbell*, Charles E Canter*, Hélène Carabin*, Jonathan Carapetis*, Loreto Carmona*, Claudia Cella*, Fiona Charlson*, Honglei Chen*, Andrew Tai-Ann Cheng*, David Chou*, Sumeet S Chugh*, Luc E Coffeng*, Steven D Colan*, Samantha Colquhoun*, K Ellicott Colson*, John Condon*, Myles D Connor*, Leslie T Cooper*, Matthew Corriere*, Monica Cortinovis*, Karen Courville de Vaccaro*, William Couser*, Benjamin C Cowie*, Michael H Criqui*, Marita Cross*, Kaustubh C Dabhadkar*, Manu Dahiya*, Nabila Dahodwala*, James Damsere-Derry*, Goodarz Danaei*, Adrian Davis*, Diego De Leo*, Louisa Degenhardt*, Robert Dellavalle*, Allyne Delossantos*, Julie Denenberg*, Sarah Derrett*, Don C Des Jarlais*, Samath D Dharmaratne*, Mukesh Dherani*, Cesar Diaz-Torne*, Helen Dolk*, E Ray Dorsey*, Tim Driscoll*, Herbert Duber*, Beth Ebel*, Karen Edmond*, Alexis Elbaz*, Suad Eltahir Ali*, Holly Erskine*, Patricia J Erwin*, Patricia Espindola*, Stalin E Ewoigbokhan*, Farshad Farzadfar*, Valery Feigin*, David T Felson*, Alize Ferrari*, Cleusa P Ferri*, Eric M Fèvre*, Mariel M Finucane*, Seth Flaxman*, Louise Flood*, Kyle Foreman*, Mohammad H Forouzanfar*, Francis Gerry R Fowkes*, Marlene Fransen*, Michael K Freeman*, Belinda J Gabbe*, Sherine E Gabriel*, Emmanuela Gakidou*, Hammad A Ganatra*, Bianca Garcia*, Flavio Gaspari*, Richard F Gillum*, Gerhard Gmel*, Diego Gonzalez-Medina*, Richard Gosselin*, Rebecca Grainger*, Bridget Grant*, Justina Groeger*, Francis Guillemin*, David Gunnell*, Ramyani Gupta*, Juanita Haagsma*, Holly Hagan*, Yara A Halasa*, Wayne Hall*, Diana Haring*, Josep Maria Haro*, James E Harrison*, Rasmus Havmoeller*, Roderick J Hay*, Hideki Higashi*, Catherine Hill*, Bruno Hoen*, Howard Hoffman*, Peter J Hotez*, Damian Hoy*, John J Huang*, Sydney E Ibeanusi*, Kathryn H Jacobsen*, Spencer L James*, Deborah Jarvis*, Rashmi Jasrasaria*, Sudha Jayaraman*, Nicole Johns*, Jost B Jonas*, Ganesan Karthikeyan*, Nicholas Kassebaum*, Norito Kawakami*, Andre Keren*, Jon-Paul Khoo*, Charles H King*, Lisa Marie Knowlton*, Olive Kobusingye*, Adofo Koranteng*, Rita Krishnamurthi*, Francine Laden*, Ratilal Lalloo*, Laura L Laslett*, Tim Lathlean*, Janet L Leasher*, Yong Yi Lee*, James Leigh*, Daphna Levinson*, Stephen S Lim*, Elizabeth Limb*, John Kent Lin*, Michael Lipnick*, Steven E Lipshultz*, Wei Liu*, Maria Loane*, Summer Lockett Ohno*, Ronan Lyons*, Jacqueline Mabweijano*, Michael F MacIntyre*, Reza Malekzadeh*, Leslie Mallinger*, Sivabalan Manivannan*, Wagner Marcenes*, Lyn March*, David J Margolis*, Guy B Marks*, Robin Marks*, Akira Matsumori*, Richard Matzopoulos*, Bongani M Mayosi*, John H McAnulty*, Mary M McDermott*, Neil McGill*, John McGrath*, Maria Elena Medina-Mora*, Michele Meltzer*, Ziad A Memish*, George A Mensah*, Tony R Merriman*, Ana-Claire Meyer*, Valeria Miglioli*, Matthew Miller*, Ted R Miller*, Philip B Mitchell*, Charles Mock*, Ana Olga Mocumbi*, Terrie E Moffitt*, Ali A Mokdad*, Lorenzo Monasta*, Marcella Montico*, Maziar Moradi-Lakeh*, Andrew Moran*, Lidia Morawska*, Rintaro Mori*, Michele E Murdoch*, Michael K Mwaniki*, Kovin Naidoo*, M Nathan Nair*, Luigi Naldi*, K M Venkat Narayan*, Paul K Nelson*, Robert G Nelson*, Michael C Nevitt*, Charles R Newton*, Sandra Nolte*, Paul Norman*, Rosana Norman*, Martin O'Donnell*, Simon O'Hanlon*, Casey Olives*, Saad B Omer*, Katrina Ortblad*, Richard Osborne*, Doruk Ozgediz*, Andrew Page*, Bishnu Pahari*, Jeyaraj Durai Pandian*, Andrea Panozo Rivero*, Scott B Patten*, Neil Pearce*, Rogelio Perez Padilla*, Fernando Perez-Ruiz*, Norberto Perico*, Konrad Pesudovs*, David Phillips*, Michael R Phillips*, Kelsey Pierce*, Sébastien Pion*, Guilherme V Polanczyk*, Suzanne Polinder*, C Arden Pope III*, Svetlana Popova*, Esteban Porrini*, Farshad Pourmalek*, Martin Prince*, Rachel L Pullan*, Kapa D Ramaiah*, Dharani Ranganathan*, Homie Razavi*, Mathilda Regan*, Jürgen T Rehm*, David B Rein*, Guiseppe Remuzzi*, Kathryn Richardson*, Frederick P Rivara*, Thomas Roberts*, Carolyn Robinson*, Felipe Rodriguez De Leòn*, Luca Ronfani*, Robin Room*, Lisa C Rosenfeld*, Lesley Rushton*, Ralph L Sacco*, Sukanta Saha*, Uchechukwu Sampson*, Lidia Sanchez-Riera*, Ella Sanman*, David C Schwebel*, James Graham Scott*, Maria Segui-Gomez*, Saeid Shahraz*, Donald S Shepard*, Hwashin Shin*, Rupak Shivakoti*, Donald Silberberg*, David Singh*, Gitanjali M Singh*, Jasvinder A Singh*, Jessica Singleton*, David A Sleet*, Karen Sliwa*, Emma Smith*, Jennifer L Smith*, Nicolas J C Stapelberg*, Andrew Steer*, Timothy Steiner*, Wilma A Stolk*, Lars Jacob Stovner*, Christopher Sudfeld*, Sana Syed*, Giorgio Tamburlini*, Mohammad Tavakkoli*, Hugh R Taylor*, Jennifer A Taylor*, William J Taylor*, Bernadette Thomas*, W Murray Thomson*, George D Thurston*, Imad M Tleyjeh*, Marcello Tonelli*, Jeffrey A Towbin*, Thomas Truelsen*, Miltiadis K Tsilimbaris*, Clotilde Ubeda*, Eduardo A Undurraga*, Marieke J van der Werf*, Jim van Os*, Monica S Vavilala*, N Venketasubramanian*, Mengru Wang*, Wenzhi Wang*, Kerrianne Watt*, David J Weatherall*, Martin A Weinstock*, Robert Weintraub*, Marc G Weisskopf*, Myrna M Weissman*, Richard A White*, Harvey Whiteford*, Natasha Wiebe*, Steven T Wiersma*, James D Wilkinson*, Hywel C Williams*, Sean R M Williams*, Emma Witt*, Frederick Wolfe*, Anthony D Woolf*, Sarah Wulf*, Pon-Hsiu Yeh*, Anita K M Zaidi*, Zhi-Jie Zheng*, David Zonies*, Alan D Lopez†

Summary

Background Measuring disease and injury burden in populations requires a composite metric that captures both premature mortality and the prevalence and severity of ill-health. The 1990 Global Burden of Disease study proposed disability-adjusted life years (DALYs) to measure disease burden. No comprehensive update of disease burden worldwide incorporating a systematic reassessment of disease and injury-specific epidemiology has been done since the 1990 study. We aimed to calculate disease burden worldwide and for 21 regions for 1990, 2005, and 2010 with methods to enable meaningful comparisons over time.

Methods We calculated DALYs as the sum of years of life lost (YLLs) and years lived with disability (YLDs). DALYs were calculated for 291 causes, 20 age groups, both sexes, and for 187 countries, and aggregated to regional and global estimates of disease burden for three points in time with strictly comparable definitions and methods. YLLs were calculated from age-sex-country-time-specific estimates of mortality by cause, with death by standardised lost

Lancet 2012; 380: 2197–223

This online publication has been corrected. The first corrected version appeared at thelancet.com on February 22, 2013, the second corrected version appeared on August 15, 2014

See **Comment** pages 2053, 2054, 2055, 2058, 2060, 2062, and 2063

See **Special Report** page 2067

See **Articles** pages 2071, 2095, 2129, 2144, 2163, and 2224

Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010

Theo Vos, Abraham D Flaxman, Mohsen Naghavi, Rafael Lozano, Catherine Michaud, Majid Ezzati, Kenji Shibuya, Joshua A Salomon, Safa Abdalla, Victor Aboyans*, Jerry Abraham*, Ilana Ackerman*, Rakesh Aggarwal*, Stephanie Y Ahn*, Mohammed K Ali*, Mohammad A AlMazroa*, Miriam Alvarado*, H Ross Anderson*, Laurie M Anderson*, Kathryn G Andrews*, Charles Atkinson*, Larry M Baddour*, Adil N Bahalim*, Suzanne Barker-Collo*, Lope H Barrero*, David H Bartels*, Maria-Gloria Basáñez*, Amanda Baxter*, Michelle L Bell*, Emelia J Benjamin*, Derrick Bennett*, Eduardo Bernabé*, Kavi Bhalla*, Bishal Bhandari*, Boris Bikbov*, Aref Bin Abdulhak*, Gretchen Birbeck*, James A Black*, Hannah Blencowe*, Jed D Blore*, Fiona Blyth*, Ian Bolliger*, Audrey Bonaventure*, Soufiane Boufous*, Rupert Bourne*, Michel Boussinesq*, Tasanee Braithwaite*, Carol Brayne*, Lisa Bridgett*, Simon Brooker*, Peter Brooks*, Traolach S Brugha*, Claire Bryan-Hancock*, Chiara Bucello*, Rachelle Buchbinder*, Geoffrey Buckle*, Christine M Budke*, Michael Burch*, Peter Burney*, Roy Burstein*, Bianca Calabria*, Benjamin Campbell*, Charles E Canter*, Hélène Carabin*, Jonathan Carapetis*, Loreto Carmona*, Claudia Cella*, Fiona Charlson*, Honglei Chen*, Andrew Tai-Ann Cheng*, David Chou*, Sumeet S Chugh*, Luc E Coffeng*, Steven D Colan*, Samantha Colquhoun*, K Ellicott Colson*, John Condon*, Myles D Connor*, Leslie T Cooper*, Matthew Corriere*, Monica Cortinovis*, Karen Courville de Vaccaro*, William Couser*, Benjamin C Cowie*, Michael H Criqui*, Marita Cross*, Kaustubh C Dabhadkar*, Manu Dahiya*, Nabila Dahodwala*, James Damsere-Derry*, Goodarz Danaei*, Adrian Davis*, Diego De Leo*, Louisa Degenhardt*, Robert Dellavalle*, Allyne Delossantos*, Julie Denenberg*, Sarah Derrett*, Don C Des Jarlais*, Samath D Dharmaratne*, Mukesh Dherani*, Cesar Diaz-Torne*, Helen Dolk*, E Ray Dorsey*, Tim Driscoll*, Herbert Duber*, Beth Ebel*, Karen Edmond*, Alexis Elbaz*, Suad Eltahir Ali*, Holly Erskine*, Patricia J Erwin*, Patricia Espindola*, Stalin E Ewoigbokhan*, Farshad Farzadfar*, Valery Feigin*, David T Felson*, Alize Ferrari*, Cleusa P Ferri*, Eric M Fèvre*, Mariel M Finucane*, Seth Flaxman*, Louise Flood*, Kyle Foreman*, Mohammad H Forouzanfar*, Francis Gerry R Fowkes*, Richard Franklin*, Marlene Fransen*, Michael K Freeman*, Belinda J Gabbe*, Sherine E Gabriel*, Emmanuela Gakidou*, Hammad A Ganatra*, Bianca Garcia*, Flavio Gaspari*, Richard F Gillum*, Gerhard Gmel*, Richard Gosselin*, Rebecca Grainger*, Justina Groeger*, Francis Guillemin*, David Gunnell*, Ramyani Gupta*, Juanita Haagsma*, Holly Hagan*, Yara A Halasa*, Wayne Hall*, Diana Haring*, Josep Maria Haro*, James E Harrison*, Rasmus Havmoeller*, Roderick J Hay*, Hideki Higashi*, Catherine Hill*, Bruno Hoen*, Howard Hoffman*, Peter J Hotez*, Damian Hoy*, John J Huang*, Sydney E Ibeanusi*, Kathryn H Jacobsen*, Spencer L James*, Deborah Jarvis*, Rashmi Jasrasaria*, Sudha Jayaraman*, Nicole Johns*, Jost B Jonas*, Ganesan Karthikeyan*, Nicholas Kassebaum*, Norito Kawakami*, Andre Keren*, Jon-Paul Khoo*, Charles H King*, Lisa Marie Knowlton*, Olive Kobusingye*, Adofo Koranteng*, Rita Krishnamurthi*, Ratilal Lalloo*, Laura L Laslett*, Tim Lathlean*, Janet L Leasher*, Yong Yi Lee*, James Leigh*, Stephen S Lim*, Elizabeth Limb*, John Kent Lin*, Michael Lipnick*, Steven E Lipshultz*, Wei Liu*, Maria Loane*, Summer Lockett Ohno*, Ronan Lyons*, Jixiang Ma*, Jacqueline Mabweijano*, Michael F MacIntyre*, Reza Malekzadeh*, Leslie Mallinger*, Sivabalan Manivannan*, Wagner Marcenes*, Lyn March*, David J Margolis*, Guy B Marks*, Robin Marks*, Akira Matsumori*, Richard Matzopoulos*, Bongani M Mayosi*, John H McAnulty*, Mary M McDermott*, Neil McGill*, John McGrath*, Maria Elena Medina-Mora*, Michele Meltzer*, Ziad A Memish*, George A Mensah*, Tony R Merriman*, Ana-Claire Meyer*, Valeria Miglioli*, Matthew Miller*, Ted R Miller*, Philip B Mitchell*, Ana Olga Mocumbi*, Terrie E Moffitt*, Ali A Mokdad*, Lorenzo Monasta*, Marcella Montico*, Maziar Moradi-Lakeh*, Andrew Moran*, Lidia Morawska*, Rintaro Mori*, Michele E Murdoch*, Michael K Mwaniki*, Kovin Naidoo*, M Nathan Nair*, Luigi Naldi*, K M Venkat Narayan*, Paul K Nelson*, Robert G Nelson*, Michael C Nevitt*, Charles R Newton*, Sandra Nolte*, Paul Norman*, Rosana Norman*, Martin O'Donnell*, Simon O'Hanlon*, Casey Olives*, Saad B Omer*, Katrina Ortblad*, Richard Osborne*, Doruk Ozgediz*, Andrew Page*, Bishnu Pahari*, Jeyaraj Durai Pandian*, Andrea Panozo Rivero*, Scott B Patten*, Neil Pearce*, Rogelio Perez Padilla*, Fernando Perez-Ruiz*, Norberto Perico*, Konrad Pesudovs*, David Phillips*, Michael R Phillips*, Kelsey Pierce*, Sébastien Pion*, Guilherme V Polanczyk*, Suzanne Polinder*, C Arden Pope III*, Svetlana Popova*, Esteban Porrini*, Farshad Pourmalek*, Martin Prince*, Rachel L Pullan*, Kapa D Ramaiah*, Dharani Ranganathan*, Homie Razavi*, Mathilda Regan*, Jürgen T Rehm*, David B Rein*, Guiseppe Remuzzi*, Kathryn Richardson*, Frederick P Rivara*, Thomas Roberts*, Carolyn Robinson*, Felipe Rodriguez De Leòn*, Luca Ronfani*, Robin Room*, Lisa C Rosenfeld*, Lesley Rushton*, Ralph L Sacco*, Sukanta Saha*, Uchechukwu Sampson*, Lidia Sanchez-Riera*, Ella Sanman*, David C Schwebel*, James Graham Scott*, Maria Segui-Gomez*, Saeid Shahraz*, Donald S Shepard*, Hwashin Shin*, Rupak Shivakoti*, Donald Silberberg*, David Singh*, Gitanjali M Singh*, Jasvinder A Singh*, Jessica Singleton*, David A Sleet*, Karen Sliwa*, Emma Smith*, Jennifer L Smith*, Nicolas J C Stapelberg*, Andrew Steer*, Timothy Steiner*, Wilma A Stolk*, Lars Jacob Stovner*, Christopher Sudfeld*, Sana Syed*, Giorgio Tamburlini*, Mohammad Tavakkoli*, Hugh R Taylor*, Jennifer A Taylor*, William J Taylor*, Bernadette Thomas*, W Murray Thomson*, George D Thurston*, Imad M Tleyjeh*, Marcello Tonelli*, Jeffrey A Towbin*, Thomas Truelsen*, Miltiadis K Tsilimbaris*, Clotilde Ubeda*, Eduardo A Undurraga*, Marieke J van der Werf*, Jim van Os*, Monica S Vavilala*, N Venketasubramanian*, Mengru Wang*, Wenzhi Wang*, Kerrianne Watt*, David J Weatherall*, Martin A Weinstock*, Robert Weintraub*, Marc G Weisskopf*, Myrna M Weissman*, Richard A White*, Harvey Whiteford*, Steven T Wiersma*, James D Wilkinson*, Hywel C Williams*, Sean R M Williams*, Emma Witt*, Frederick Wolfe*, Anthony D Woolf*, Sarah Wulf*, Pon-Hsiu Yeh*, Anita K M Zaidi*, Zhi-Jie Zheng*, David Zonies*, Alan D Lopez†, Christopher J L Murray†‡*

Summary

Background Non-fatal health outcomes from diseases and injuries are a crucial consideration in the promotion and monitoring of individual and population health. The Global Burden of Disease (GBD) studies done in 1990 and 2000 have been the only studies to quantify non-fatal health outcomes across an exhaustive set of disorders at the global and regional level. Neither effort quantified uncertainty in prevalence or years lived with disability (YLDs).

Methods Of the 291 diseases and injuries in the GBD cause list, 289 cause disability. For 1160 sequelae of the 289 diseases and injuries, we undertook a systematic analysis of prevalence, incidence, remission, duration, and excess mortality. Sources included published studies, case notification, population-based cancer registries, other disease registries, antenatal clinic serosurveillance, hospital discharge data, ambulatory care data, household surveys, other surveys, and cohort studies. For most sequelae, we used a Bayesian meta-regression method, DisMod-MR,

Lancet 2012; 380: 2163–96

This online publication has been corrected. The corrected version first appeared at thelancet.com on February 22, 2012, the second corrected version appeared on August 15, 2014

See Comment pages 2053, 2054, 2055, 2058, 2060, 2062, and 2063

See Special Report page 2067