Baseline drivers of lymphatic filariasis in Burkina Faso

Michelle C. Stanton[1], David H. Molyneux[1], Dominique Kyelem[2†], Roland W. Bougma[3], Benjamin G. Koudou[1], Louise A. Kelly-Hope[1]

[1]*Centre for Neglected Tropical Diseases, Liverpool School of Tropical Medicine, Liverpool, United Kingdom;* [2]*Neglected Tropical Diseases Support Centre, The Task Force for Global Health, Atlanta, USA;* [3]*Ministry of Health, Ouagadougou, Burkina Faso*

Abstract. Lymphatic filariasis (LF) is a parasitic disease that is endemic throughout sub-Saharan Africa, infecting approximately 40 million people. In Burkina Faso, mass drug administration (MDA) for LF with ivermectin and albendazole has been ongoing since 2001, and by 2006 all endemic health districts were receiving MDA with a therapeutic coverage of at least 65%. As MDA activities scale down, the focus is now on targeting areas where LF transmission persists with alternative elimination strategies. This study explored the relationship between village-level, baseline LF prevalence data collected in 2000 with publicly available meteorological, environmental and demographic variables in order to determine the factors that influence the geographical distribution of the disease. A fitted multiple logistic regression model indicated that the length of the rainy season, variability in normalized difference vegetation index (NDVI) and population density were significantly positively associated with LF prevalence, whereas total annual rainfall, average June-September temperature, mean NDVI, elevation and the area of cotton crops were significantly negatively associated. This model was used to produce a baseline LF risk map for Burkina Faso. An extended model which incorporated potential socio-demographic risk factors also indicated a significant positive relationship between LF prevalence and wealth. In overlaying the baseline LF risk map with the number of MDA rounds, plus an insecticide-treated net (ITN) ownership measure, the central southern area of the country was highlighted as an area where baseline LF prevalence was high and ITN coverage relatively low (<50%), while at least 10 rounds of MDA had been undertaken, suggesting that more concentrated efforts will be needed to eliminate the disease in these areas.

Keywords: lymphatic filariasis, neglected tropical diseases, control and elimination, remote sensing, Burkina Faso.

Introduction

Lymphatic filariasis (LF), caused by the parasite *Wuchereria bancrofti*, is a disease that is endemic throughout sub-Saharan Africa. Approximately 400 million Africans are at risk of the disease, with an estimated 40 million being infected, many of which experience severe disabilities, including lymphoedema and hydroceles as reported by the World Health Organization (WHO, 2012a). LF has been targeted for elimination by 2020, with the primary method of elimination being the use of annual mass drug administration (MDA) campaigns with ivermectin and albendazole (or diethylcarbamazine (DEC) and albendazole in countries where onchocerciasis is not endemic) to interrupt transmission of the disease (WHO, 2010) for a minimum of five consecutive years, with coverage of

at least 65% (WHO, 2012b). The Global Programme to Eliminate Lymphatic Filariasis (GPELF) was launched in 2000, and the West African country of Burkina Faso was one of the first countries to conduct a national baseline LF prevalence survey at the health district level in order to estimate the burden of the disease. It resulted in an estimated 29.2% national prevalence with the highest district-level prevalence being 74% (Programme National d'Elimination de la Filariose Lymphatique Burkina Faso, 2012). Shortly following these results, Burkina Faso launched their national programme for the elimination of LF and began, in 2001, implementing MDA within four health districts. MDA then scaled up rapidly across the country, such that by 2006, therapeutic coverage levels were greater than 65% in all endemic districts (The Global Atlas of Helminth Infections, 2012).

Prior to the initiation of MDA for LF, between 1974-2002, the onchocerciasis control programme (OCP) in West Africa was operational in large parts of Burkina Faso (Boatin, 2008). Initially, vector control methods (predominantly larviciding) were applied to interrupt the transmission of the disease. From 1987, ivermectin was used in areas where larviciding had proved to be ineffective (Fig. 1) due to factors such as

Corresponding author:
Michelle C. Stanton
Centre for Neglected Tropical Diseases
Liverpool School of Tropical Medicine
Pembroke Place, Liverpool, L3 5QA, United Kingdom
Tel./Fax +44 151 705 2993
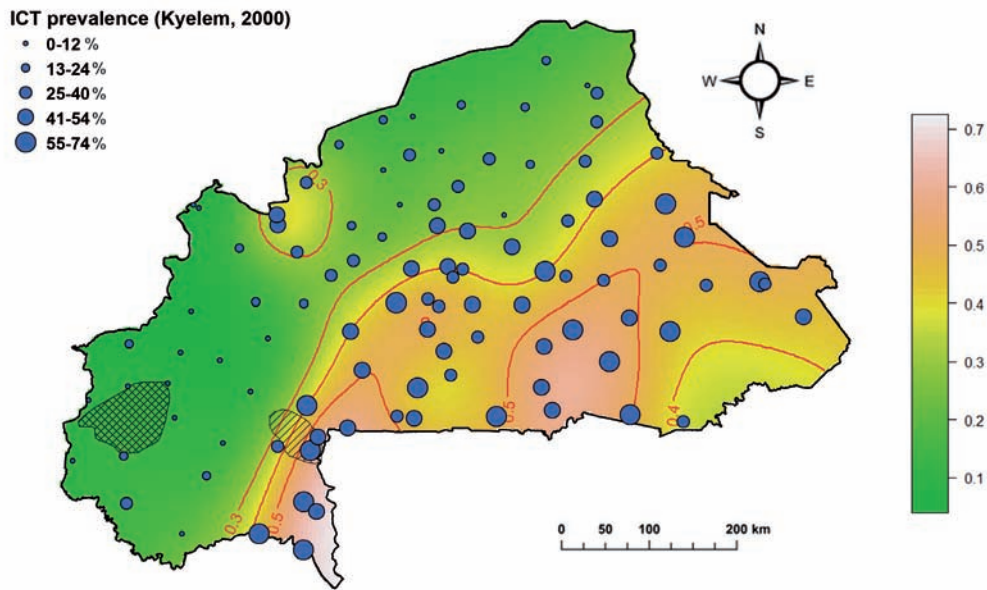E-mail: michelle.stanton@liverpool.ac.uk
† deceased

Fig. 1. Location of 102 villages in Burkina Faso surveyed in 2000 and their corresponding ICT prevalence.

poor larviciding coverage, population migration and larviciding stopping too early (Borsboom et al., 2003). These interventions impacted the prevalence of LF (Kyelem et al., 2003, 2005) and also affected the distribution of the population within Burkina Faso, as the threat of onchocerciasis had previously driven many people away from at-risk areas in river valleys. Following up to 13 years of MDA, these activities are starting to scale down in Burkina Faso. However, in order to achieve elimination targets, it is crucial that alternative control methods or longer term, annual MDA are implemented in areas where *W. bancrofti* transmission persists. In this study we consider the use of fine-scale environmental, meteorological and demographic data to produce LF risk maps based on baseline survey data. These maps are planned to highlight the areas that have conditions most suited to high LF transmission, and as such allow the national programme manager and policy makers to target areas where MDA may be prolonged, or where alternative control strategies will be needed.

Risk maps based on environmental information are particularly valuable in resource-poor settings where epidemiological data are scarce, yet meteorological and environmental data are abundant. In sub-Saharan Africa, such risk maps have been most commonly developed for malaria (Machault et al., 2011), although other examples include LF (Lindsay and Thomas, 2000), schistosomiasis (Brooker, 2007; Yang et al., 2012), leishmaniasis (Elnaiem et al., 2003) and human African trypanosomiasis (Odiit et al., 2006) as

well as other infectious diseases such as meningococcal meningitis (Molesworth et al., 2003). This type of risk map can be easily updated, e.g. when the geographical landscape, climate and/or land use changes over time, without requiring any additional epidemiological data to be collected.

With regard to LF, little research has been undertaken in exploring the environmental and meteorological characteristics that promote the transmission of LF in sub-Saharan Africa. In 1975, Brengues defined seven LF transmission zones, each being defined by rainfall (amount of rain, and length of the rainy season), temperature, elevation, forest cover and population density (Brengues, 1975). Burkina Faso spans three of these zones. Lindsay and Thomas (2000) took a more rigorous approach and developed a model to predict the presence/absence of LF throughout Africa, using rainfall and temperature as predictors and epidemiological data collected at 477 sites over the period 1960-1995. However, no estimate of the prevalence of LF was made.

Kelly-Hope et al. (2006) explored the relationship between LF and malaria using data on vegetation cover, annual average temperature, average annual precipitation and absolute humidity to assess their influence on the distribution of the parasites for both diseases in West Africa. The authors found a weak positive relationship between LF and both vegetation cover and humidity, and a weak negative relationship between LF and temperature. Further, a significant negative relationship between LF and malaria preva-

lence was observed, which the authors speculate may be attributable to the differential distribution and vector competence of the mosquito species in the study area. In addition to these environmental risk factors, wealth measures, and proxies thereof, are thought to be related to LF prevalence (Gyapong et al., 1996; Molyneux, 2008). For example, housing conditions, access to clean water and sanitation facilities may provide conditions that are conducive to the survival of the LF vector (Erlanger et al., 2005; Gazzinelli et al., 2012; Upadhyayula et al., 2012). Further, education status has been shown to be related to LF prevalence in India (Upadhyayula et al., 2012), and personal protection measures (use of bed nets, mosquito repellents and insecticides) have been related to LF in Kenya (Mwobobia and Mitsui, 1999).

The objective of this study was to explore relationship between LF prevalence prior to the interventions implemented by the national LF control programme and potential meteorological, environmental and socio-demographic drivers, using publicly available resources. This information will be used to highlight areas where LF is likely to persist despite repeated rounds of MDA and the recent scaling up of vector control strategies (Institut National de la Statistique et de la Démographie, 2012).

Materials and methods

Data sources

Epidemiological data

Baseline LF prevalence data were collected in 2000 (Gyapong et al., 2002; Kyelem, 2007). Prevalence surveys were undertaken in 102 randomly selected villages across the 53 health districts of Burkina Faso using the immunochromatographic test (ICT) for circulating filarial antigen. Sampling methods were based on the recommendations of the RAGFIL study, such that sampled villages were a minimum of 50 km apart (Gyapong and Remme, 2001). Within each village, 50-100 adults (≥15 years) were tested, with equal numbers of males and females. The geographical coordinates of each of the sampled villages were also recorded.

Gridded meteorological, environmental and demographic data

Gridded meteorological, environmental and demographic data were extracted from publicly available sources at the finest spatial resolution that could be obtained. Where possible, data were extracted over a

Table 1. Meteorological, environmental and demographic metadata used for risk profiling of lymphatic filariasis in Burkina Faso.

Data	Resolution	Time period	Source	Website*
Rainfall	0.1° by 0.1°	1996-2005	Famine Early Warning System (FEWS)	1
June-September temperature	0.1° by 0.1°	1996-2005	MODIS	2
NDVI	250 x 250 m	August 2000-July 2001	LandDAAC MODIS	3
Soil moisture	0.5° by 0.5°	1996-2005	NOAA NCEP CPC	4
Tree cover	0.5 x 0.5 km	2001	MODIS VCF	5
Land cover	1 x 1 km	2000	JRC Global Land Cover Project	6
Crop area and cotton area	10 x 10 km	2000	Spatial Production Allocation Model (SPAM)	7
Elevation	1 x 1 km	NA	GTOPO30	8
Rivers	NA	2000	UNEP/FAO	9
Population density	1 km by 1 km	2000	UNEP	10
Census variables[a]	Health district	1996	Institut National de la Statistique de la Démographie	11
DHS variables[b]	NA	1998	Demographic and Health Survey	12

[a]Literacy, level of education, economically active population, infant mortality, access to drinkable water, sanitation; [b]wealth index, water and sanitation access.
*1) http://iridl.ldeo.columbia.edu; 2) http://neo.sci.gsfc.nasa.gov; 3) http://iridl.ldeo.columbia.edu; 4) http://iridl.ldeo.columbia.edu; 5) http://glcf.umiacs.umd.edu/data/vcf/; 6) http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php; 7) http://harvestchoice.org/products/data; 8) http://eros.usgs.gov/; 9) http://www.fao.org/geonetwork/srv/en/main.home; 10) http://na.unep.net/siouxfalls/datasets/datalist.php; 11) http://www.insd.bf/fr/; 12) http://www.measuredhs.com/.

10-year time period (1996-2005) and averages were taken over this time period to account for the possibility of anomalies in the data (Table 1). From these data, average total annual rainfall, average number of wet days in a year, mean normalized difference vegetation index (NDVI), the standard deviation (SD) of NDVI, average soil moisture, percentage of tree cover, area being harvested for any crop, area being harvested for cotton, distance to the nearest river and population density values were derived and extracted for each of the surveyed villages. Summaries of each of these variables (minimum, mean, maximum) were then calculated and tabulated.

Province-level socio-demographic data

Demographic variables relating to education (the literate proportion of the population, proportion of the population who had completed secondary education), wealth/health (the economically active proportion of people over 10, infant mortality during the first year of life, proportion of households without access to drinkable water, proportion of households without access to a latrine), were extracted from the 1996 national census data at the province level (Institut National de la Statistique et de la Démographie, 1996). Health district-level data were not available.

Village-level socio-demographic data

Additional demographic data were extracted from the 1998-1999 Demographic and Health Survey (DHS). Data were calculated from 4,812 households across 208 clusters in Burkina Faso (Institut National de la Statistique et de la Démographie, 2000) on numerous topics relating to the demographics and health of the population. The DHS have further developed a household-level wealth index, incorporating information of household ownership of assets, material used for housing construction, water access and sanitation facilities. A cluster-level measure of poverty was calculated, defined as the proportion of surveyed households categorised to be in either the first or second wealth quintile i.e. were amongst the poorest of the surveyed population. The poverty measure of the DHS cluster nearest to each village surveyed for LF was assigned to that village. If the nearest DHS cluster was more than 30 km from the surveyed villages, the poverty score was classed as missing. We further extracted several of the components of the wealth index that were available for this period (proportion without access to clean drinking water, proportion

without access to sanitation facilities) and obtained the proportion of people surveyed in each cluster, who had not received any form of education.

Exploratory data analysis

Spatial trends in the epidemiological data were examined by producing a map of smoothed ICT prevalence, using a smoothing spline (Hastie and Tibshirani, 1986; Wood, 2006). To explore whether any of the spatial variation in ICT prevalence within Burkina Faso could be explained by any of the meteorological, environmental or demographic variables described above, plots of each of the quantitative variables against ICT prevalence were produced and the Spearman's rank correlation coefficient calculated. Further to considering Burkina Faso as a whole, the correlation was also calculated for villages within both high and low LF prevalence zones separately, as determined by the smoothed prevalence map. As ivermectin was known to have been distributed in both around the Bourgouriba and the Dienkoa River basins under OCP (Fig. 1) (Borsboom et al., 2003; Kyelem et al., 2003, 2005), villages within these areas were excluded from the analysis.

Logistic regression analysis

A multiple logistic regression model was fitted to the prevalence data. As the goal was to produce an LF risk map for the entire country, initially only variables that could be calculated for any given location within Burkina Faso were considered for inclusion in this baseline risk model. An exhaustive screening model selection approach was applied to select which variables to include in the model. The corrected Akaike's information criterion (AIC) was used to to determine which of the possible models resulted in the best fit, and this model was then assessed by calculating the root mean square error (RMSE) of the predictions using a leave-one-out, cross-validation approach and undertaking the goodness-of-fit test of Hosmer and Lemesbow (1980). The predictive ability of the model was further assessed by calculating the sensitivity and specificity of the predictions with respect to an arbitrary threshold, here the median prevalence value. The variogram (a plot of semi-variance against the distance between two points) for the residuals of the fitted model was then estimated to determine whether or not any spatial structure remained that had not been explained by the selected variables (Cressie, 1993). This model selection and evaluation approach was
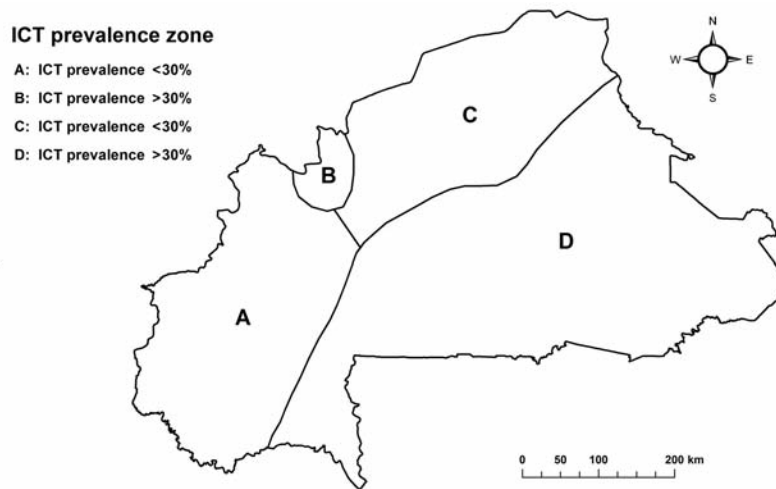
Fig. 2. High and low LF prevalence zones in Burkina Faso as defined by smoothed ICT prevalence.

then repeated with the inclusion of the DHS-derived variables, which we refer to as the socio-demographic model.

Using the results of the baseline risk model, an LF risk map was finally produced for the entire study area in order to highlight areas where LF may persist/resurge once MDA has been scaled down. This risk map was compared with the number of rounds of MDA received by each health district up to and including 2012 (Programme National d'Elimination de la Filariose Lymphatique Burkina Faso, 2012; The Global Atlas of Helminth Infections, 2012) and the regional-level insecticide-treated net (ITN) coverage as estimated by the 2010 DHS survey (Institut National de la Statistique et de la Démographie, 2012).

Results

Exploratory data analysis

Fig. 1 presents a map of the smoothed ICT prevalence, plus locations of the 102 surveyed villages in Burkina Faso. This map shows a clear spatial trend in prevalence, such that it is low in the northern and western parts of the country and high in the south-eastern half of the country. Using these observations, the country was divided into high- and low-prevalence zones, using the 0.3 contour as a threshold. As the high-prevalence areas were not contiguous, two high-prevalence zones were created. Further, due to clear environmental differences between the western low-prevalence area and the northern low prevalence area which extends into the Sahara

Table 2. Summaries of prevalence and gridded meteorological, environmental and demographic variables by zone.

Variable	Zone A (n = 21)			Zone B (n = 4)			Zone C (n = 23)			Zone D (n = 54)		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Prevalence	10	0	30	39	28	50	23	6	42	48	26	74
Rainfall	865	647	983	701	638	763	612	443	795	766	484	963
Wet days	92	73	107	74	70	78	69	54	84	87	60	114
Temperature	27	23	32	31	30	32	34	31	36	29	24	35
Mean NDVI	0.40	0.29	0.53	0.29	0.27	0.33	0.24	0.11	0.38	0.33	0.19	0.48
SD (NDVI)	0.14	0.11	0.18	0.14	0.12	0.16	0.10	0.05	0.15	0.13	0.06	0.20
Soil moisture	253	180	291	192	184	217	167	109	216	225	139	367
Tree cover	6.81	1	21	1.5	1	3	0.61	0	2	3.33	0	13
Crop area	801	0	2,239	1,539	1,146	2,200	2,154	1,162	3,115	1,488	0	4,436
Cotton area	236	0	850	17	7	26	18	0	179	59	0	1,289
Elevation	346	259	433	265	260	271	307	261	355	293	162	415
Distance to river	9.9	1.2	19.6	4.4	0.8	8.8	7.0	0	22.3	8.4	0	31.7
Population density	49	8	273	44	20	80	51	5	127	57	10	199

SD, Standard deviation

Table 3. Summary of DHS-derived variables by zone.

Variable	Zone A (n = 15)			Zone B (n = 4)			Zone C (n = 21)			Zone D (n = 42)		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
1st and 2nd wealth quintile	28	0	60	33	24	39	37	0	82	46	0	92
No education	63	28	77	66	64	67	69	36	80	65	18	82
Unsafe drinking water	4	0	27	0	0	0	3	0	29	8	0	100
No toilet facilities	69	0	100	91	89	92	80	0	100	83	0	100

desert, these were also subdivided into two zones. Thus, four zones were defined and delineated (Fig. 2).

Table 2 presents summaries of each of the gridded meteorological, environmental and demographic variables being considered for all of the 102 villages by zone. Note that only four villages are contained within zone B. Further, Table 3 presents summaries of each of the DHS-derived demographic variables by zone. Maps of province-level variables are presented in Fig. 3. As anticipated, the northern part of Burkina Faso (zones B and C) are hotter and drier than the southern part of the country (zones A and D) with a minimum June-September temperature of 30 °C and maximum total annual rainfall of 795 mm compared to minimum of 23 °C and maximum of 983 mm in the southern part. Zones A and D differ with regards to vegetation and tree cover, such that the western part of the country tends to contain more densely vegetated areas than the south-eastern part, i.e. mean NDVI and tree

cover in zone A were 0.40 (range 0.29-0.53) and 6.8% (range 1-21%) respectively compared to 0.33 (range 0.19-0.48) and 3.3% (range 0-13%) in zone D. With regards to crops, zone A tends to contain less cropland than the remainder of the country overall, however almost all of Burkina Faso's cotton crops are grown in this area. Overall, the elevation of Burkina Faso is low, however there is variability across the country with the western area (zone A) being the most highly elevated (mean 346 m, range 259-433 m).

With regard to census-derived socio-demographic variables, the western part of Burkina Faso tends to be the more prosperous part of the country, with a greater proportion of its population receiving some level of education, and having greater access to clean water and sanitation facilities (Fig. 3). Scatterplots of DHS-derived village-level socio-demographic variables (Fig. 4) partially support this observation with villages in Zone A being more wealthy. However, geo-
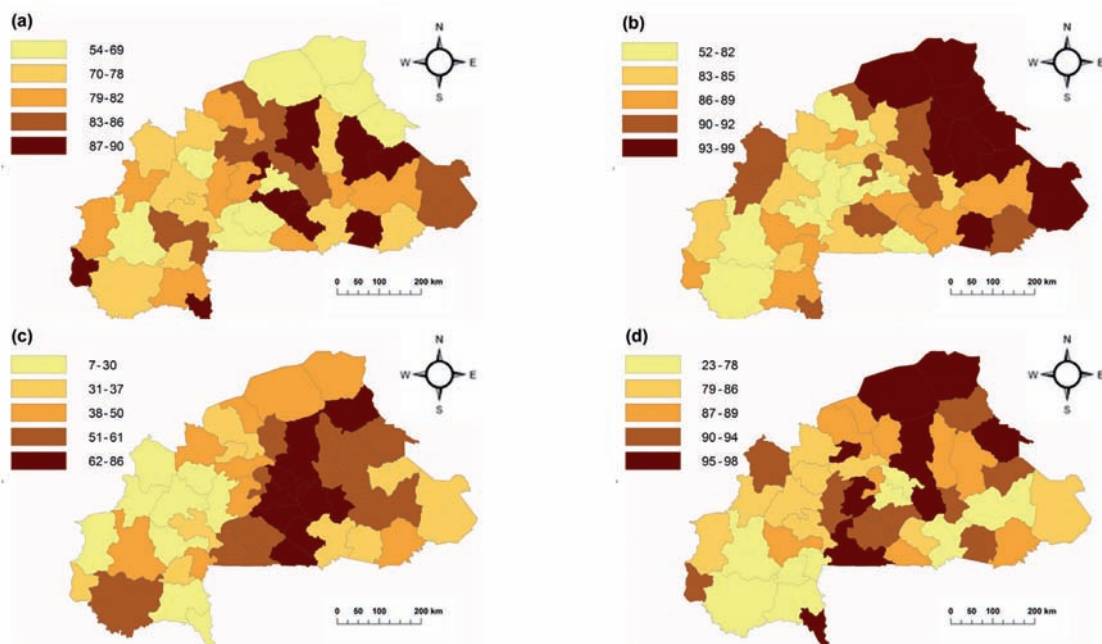

Fig. 3. Maps of 1996 census-derived province-level data: (a) proportion of people over 10 years old who are economically active; (b) proportion of people who have not completed primary level education; (c) proportion of people who do not have access to clean drinking water; and (d) proportion of people who do not have access to a latrine.
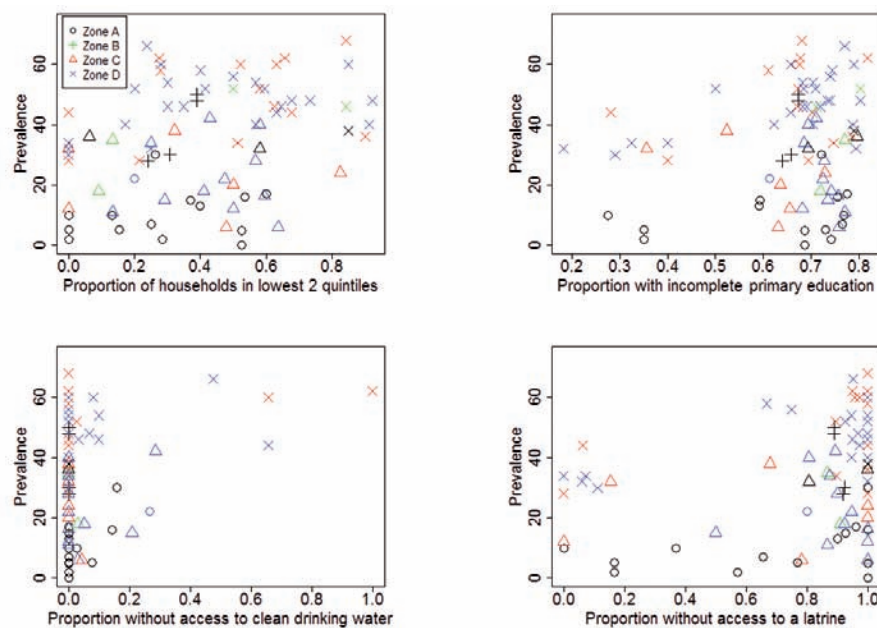
Fig. 4. Socio-demographic variables derived using the 1998-1999 DHS data from Burkina Faso, plotted against ICT prevalence.

graphical trends in education, access to clean drinking water and access to sanitation facilities are less clear.

Considering the country as a whole, the relationship between ICT prevalence and each of the gridded meteorological, environmental and demographic variables under consideration appeared weak, with the exception of elevation (Figs. 5 and 6). However, on examining the scatterplots by zone the relationships between prevalence and many of the variables appeared to follow a different trend in zone A to that of the remainder of the country. This was most apparent when considering rainfall, temperature, NDVI and, particularly, the soil moisture variables. To quantify this, Spearman's rank correlation coefficient was calculated, first using data from 95 villages (excluding seven villages within 10 km of the two OCP ivermectin distribution zones) and then separating villages in zone A and zones B-D (Table 4). Taking into account that there are only 17 villages in zone A compared to 78 in zones B-D, there still appeared to be evidence that the relationship between ICT prevalence and potential predictors in zone A do not appear to follow the same trend as those observed within zones B-D. In particular, a strong positive relationship was observed (ρ >0.5; P <0.001) between total annual rainfall, number of wet days, mean NDVI and soil moisture and ICT prevalence, and a strong negative relationship between temperature and ICT prevalence (ρ = -0.62; P <0.001) in zones B-D, whereas the relationships observed in zone A, with the exception of elevation (ρ = -0.70; P <0.002), were very weak.

With regard to cropland, zone A exhibited a positive, borderline significant, relationship between crop area and ICT prevalence (ρ = 0.41; P = 0.066), whereas the rest of the country exhibited a significant negative relationship (ρ = -0.29; P = 0.008). Due to the lack of variability in the cotton crop area measure (65/81 villages had less than 10 ha of cotton crops in their surrounding area), it was not appropriate to calculate the correlation coefficient associated with this variable for zones B-D. In zone A, the primary cotton growing area of Burkina Faso (Fig. 6), there was a weak negative relationship between cotton crop area and ICT prevalence (ρ = -0.22; P = 0.341).

As province boundaries extend across zones, the correlation between census variables and the whole country only were calculated. Of the variables considered, the proportion of economically active people over 10 (ρ = 0.23; P = 0.027), infant mortality during the first year of life (ρ = 0.24; P = 0.016) and proportion of households without access to drinkable water (ρ = 0.23; P = 0.026) were significantly correlated with LF prevalence. It was however noted that the direction of the relationship between economic activity and LF prevalence was counter-intuitive.

With regard to the DHS-derived demographic variables (Table 5), the relationship between prevalence and both wealth and access to sanitation was consistent across the country (ρ = 0.26; P = 0.023; ρ = 0.27; P = 0.018 respectively). There was also evidence of a relationship between prevalence and lack of access to clean water in zone A (ρ = 0.34; P = 0.068).
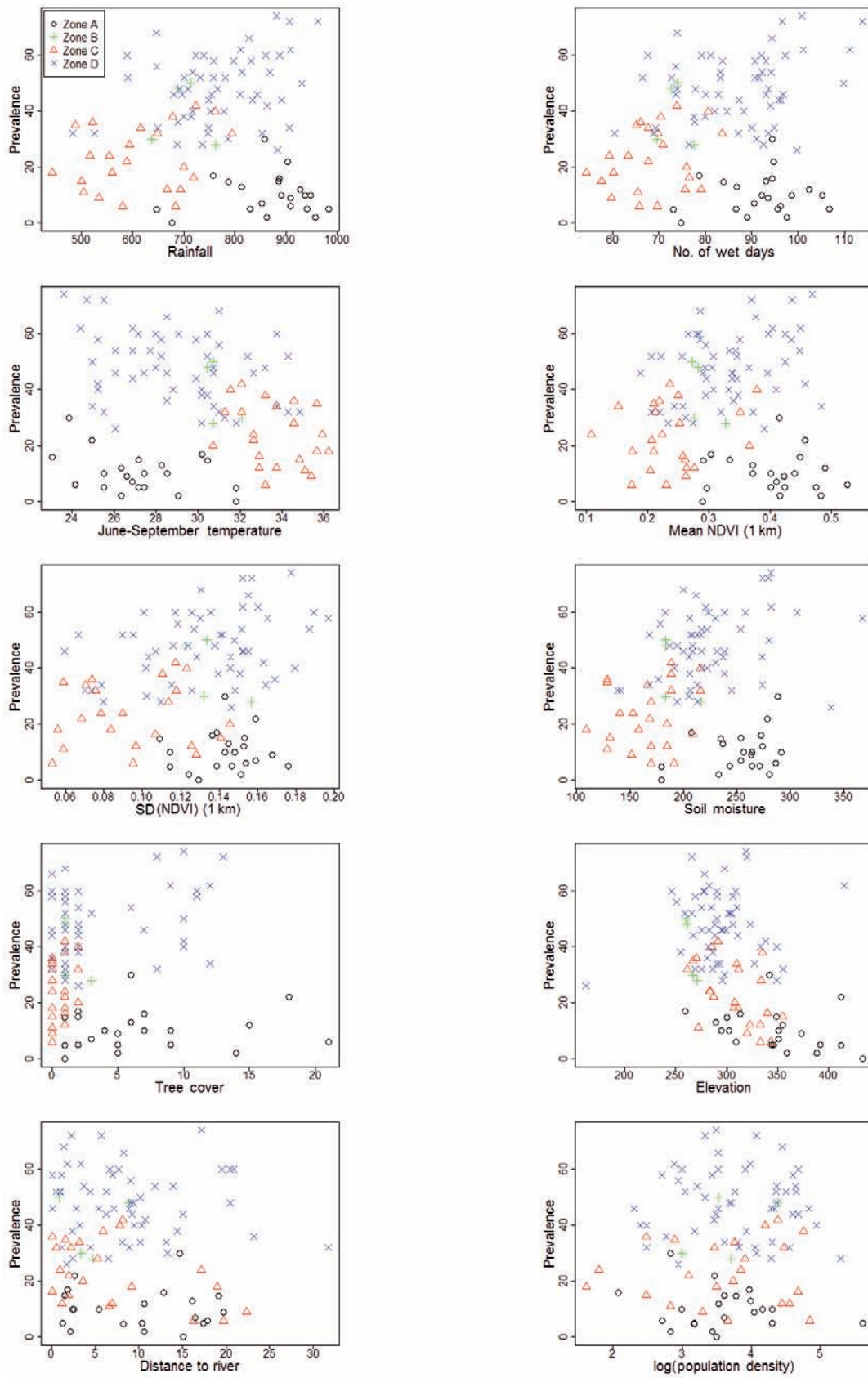
Fig. 5. Scatterplots of meteorological (rainfall, no. of wet days, June-September temperature), environmental (mean and standard deviation of NDVI, soil moisture, tree cover, elevation, distance to river) and demographic (population density) variables against ICT prevalence by zone.
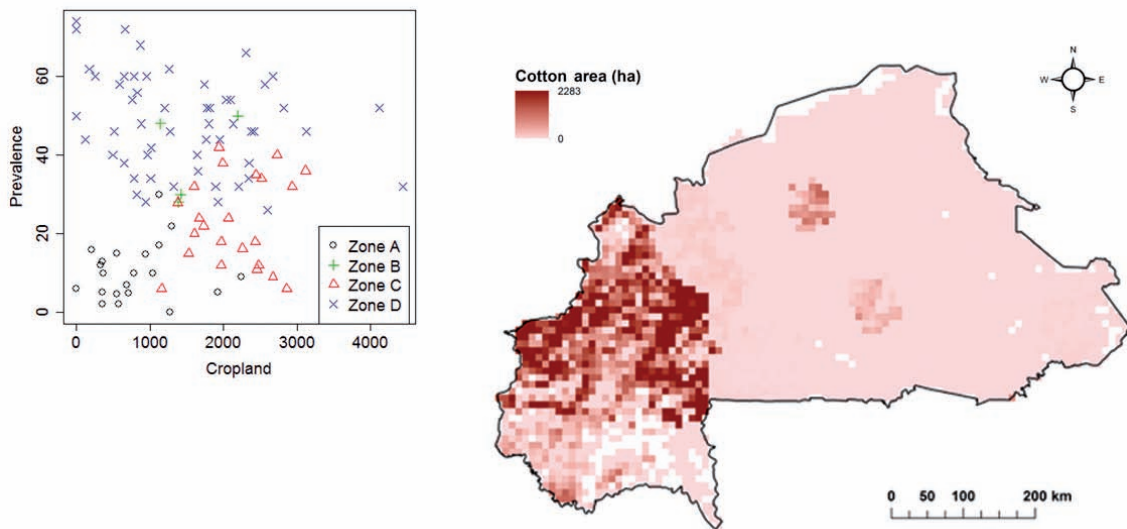
Fig. 6. Area of croplands against ICT prevalence (left) and map of area of cotton crops, measured in hectares per 5 km².

Results from logistic regression analysis

The baseline risk model

The selected baseline risk model included June-September temperature, total annual rainfall, mean number of wet days, mean and SD of NDVI, elevation, population density and cotton crop area. On undertaking model diagnostic, it was noted that there were three outliers that had a large influence on the fit of the model. Two of these villages were within 10 km of the border with Ghana, whereas one was 15 km from the more eastern OCP area. Data from these villages were excluded, and the model was refitted. The estimated coefficients from this model and their associated 95% confidence intervals are presented in Table 6.

A positive relationship was observed between LF prevalence and the number of wet days in a year, the SD of NDVI and population density, whereas a negative relationship was observed between LF prevalence and total annual rainfall, June-September temperature, mean NDVI, elevation and surrounding cotton crop area.

Observed prevalence values were then plotted against predicted values (Fig. 7), and a prediction error of 0.133 was calculated. These both indicated that although the general trend of the prevalence is accurately captured by the model, there is a lot of extra variability in the prevalence values that the model does not explain. The Hosmer and Lemeshow's goodness-of-fit test confirmed the overall poor fit (P <0.01). However, as the general trend in LF preva-

Table 4. Spearman's rank correlation between ICT prevalence and potential gridded predictor variables.

Variable	All villages		Zone A		Zone B	
	ρ	P-value	ρ	P-value	ρ	P-value
Rainfall	0.12	0.266	-0.01	0.963	0.54	<0.001
Wet days	0.22	0.032	0.18	0.483	0.55	<0.001
Temperature	-0.23	0.026	-0.31	0.229	-0.62	<0.001
Mean NDVI	0.10	0.330	-0.08	0.747	0.52	<0.001
SD (NDVI)	0.20	0.049	-0.06	0.811	0.48	<0.001
Soil moisture	0.16	0.132	0.25	0.342	0.56	<0.001
Tree cover	0.04	0.720	-0.04	0.875	0.39	<0.001
Crop area	0.02	0.854	0.41	0.066	-0.29	0.008
Elevation	-0.42	<0.001	-0.70	0.002	-0.22	0.053
Distance to river	-0.14	0.183	-0.13	0.616	0.02	0.863
Population density	0.16	0.131	0.06	0.807	0.09	0.429

Table 5. Spearman's rank correlation between ICT prevalence and potential DHS-derived predictor variables.

Variable	All villages		Zone A		Zone B	
	ρ	P-value	ρ	P-value	ρ	P-value
1st and 2nd wealth quintile	0.26	0.023	0.27	0.423	0.26	0.035
No education	0.02	0.890	0.34	0.303	0.04	0.773
Unsafe drinking water	0.09	0.423	0.57	0.068	0.12	0.331
No toilet facilities	0.27	0.018	0.34	0.306	0.28	0.023

lence has been captured by the model, it can still be used to identify areas of high/low prevalence, with the sensitivity and specificity of the model with respect to predicting whether or not prevalence is greater than the median value of 40% were 72% and 80%, respectively.

The socio-demographic model

Following the exclusion of villages that did not have a DHS cluster within a 30 km radius, villages that were within the OCP zones, and outliers, the socio-demographic model was fitted to data from 74 villages. The selected model was very similar to the selected baseline risk model, with NDVI measures being replaced by "tree cover", plus the inclusion of the poverty variable. A negative relationship was observed between tree cover and LF prevalence, whereas a positive relationship was observed between

Table 6. Estimated coefficients and their associated 95% confidence intervals on the logit-transformed scale.

Variable	log(OR)	95% CI
Model without DHS variables		
Rainfall (mm)	-0.005	-0.007, -0.003
Wet days	0.046	0.031, 0.060
Temperature (°C)	-0.155	-0.205, -0.106
Mean NDVI	-6.392	-8.296, -4.501
SD (NDVI)	11.474	8.218, 14.745
Elevation (m)	-0.015	-0.017, -0.013
Cotton area (ha)	-0.003	-0.003, -0.002
log (population density)	0.220	0.124, 0.317
Model with DHS variables		
Rainfall (mm)	-0.004	-0.006, -0.003
Wet days	0.036	0.018, 0.053
Temperature (°C)	-0.161	-0.216, -0106
Tree cover	-0.103	-0.144, -0.062
Elevation (m)	-0.015	-0.018, -0.013
Cotton area (ha)	-0.003	-0.004, -0.002
log (population density)	0.273	0.171, 0.376
Wealth	0.596	0.294, 0.900

poverty and prevalence. Fig. 7 presents a scatterplot of observations for this model *vis-à-vis* the predicted prevalence. The prediction error was 0.128, and the sensitivity and specificity were 0.73 and 0.92, respectively, indicating that this model had a better fit to the data than that obtained excluding the poverty variable. In order to further explore whether there were significant differences in prevalence at the health district-level level, an additional component in the form of a health district-level (spatially unstructured) random intercept. The inclusion of random effects into the model, however, did not improve the fit, indicating that remaining fluctuations in prevalence operated at the sub-health district-level.

The baseline risk map

The LF baseline risk map for Burkina Faso is presented in Fig. 8. A variogram fitted to the residuals from this fitted model (not shown) implied that the unexplained variability in the data was not spatially structured. These results indicated that all spatial correlation for spatial ranges of 50 km or more may be explained by environmental similarities. Due to the sampling method used, it was not possible to assess whether spatial correlation exists at shorter distances.

The baseline risk map highlight elevated levels of LF risk in the southern central and eastern part of Burkina Faso close to the border with Ghana. Further high risk areas are seen in the western and northern parts of the country, although these areas are more fragmentary. Fig. 9 presents the number of rounds of MDA received by each health district by 2012 and indicates also which districts have stopped, or are stopping, MDA at this time point following the successful reduction in LF prevalence to below 1%. Areas in the southern central and eastern part of Burkina Faso, which had high baseline prevalence, have still not managed to reduce prevalence to below 1%, despite 10-12 rounds of MDA. In areas where MDA had ceased, baseline LF risk was generally low with fragmented areas of elevated risk.
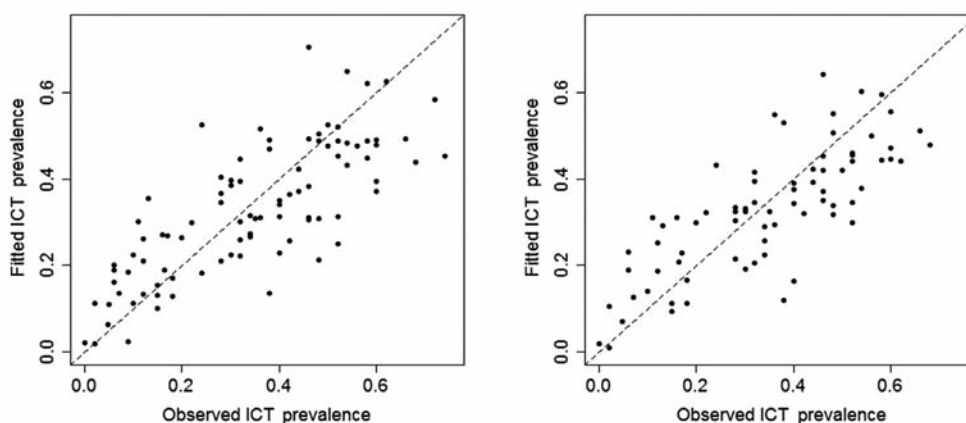
Fig. 7. Results from the fitted models: observed *versus* fitted prevalence for the logisitic regression model excluding (left) and including (right) the DHS wealth-derived estimate.

Further, regional-level ITN ownership data was extracted from the DHS survey conducted in 2010 (Institut National de la Statistique et de la Démographie, 2012). From this data it was possible to obtain a regional-level estimate of ITN coverage, defined as the proportion of surveyed households owning at least one ITN. It was observed that the southern central areas, corresponding to the regions of Central-South and Central-East, had both baseline ICT prevalence greater than 45%, and low subsequent ITN coverage in comparison to the rest of the country, i.e. 48.6% and 45.5%, respectively (Fig. 10).

Discussion

This analysis indicates that the amount and duration of rainfall, average June-September temperature, the amount and variability of vegetation, elevation, the area of cotton cultivation and population density are negatively associated with LF prevalence in Burkina Faso. These relationships make sense intuitively and may be related to the vector competence of the local *Anopheles* species and/or their specific ecological niches. Although the role of temperature, vegetation and rainfall is obvious, this is not the case for elevation. In
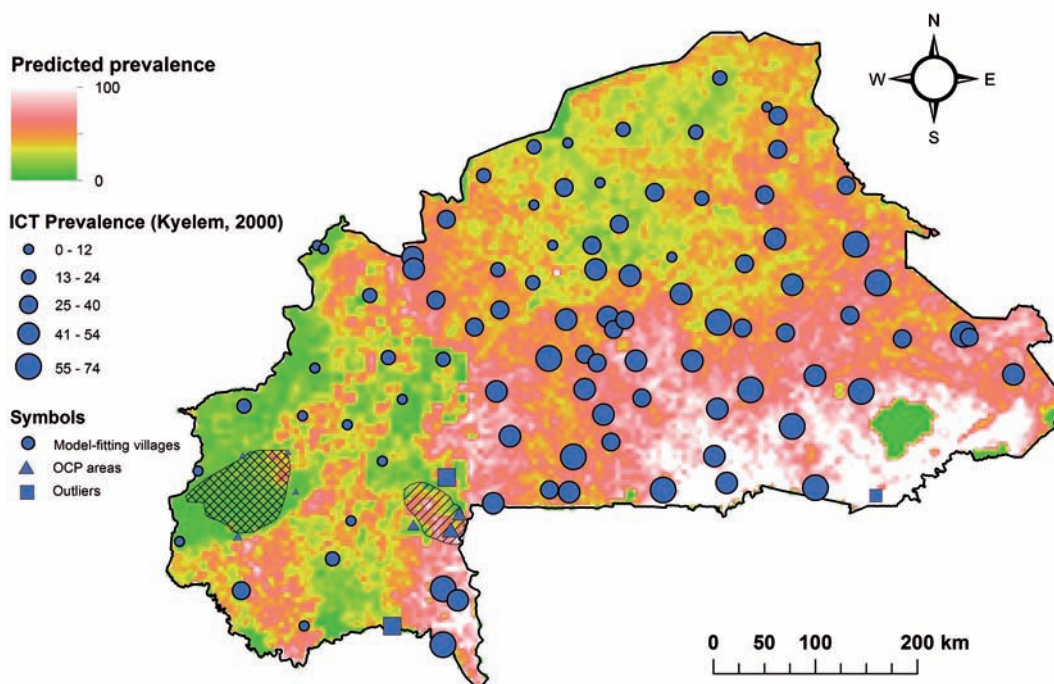

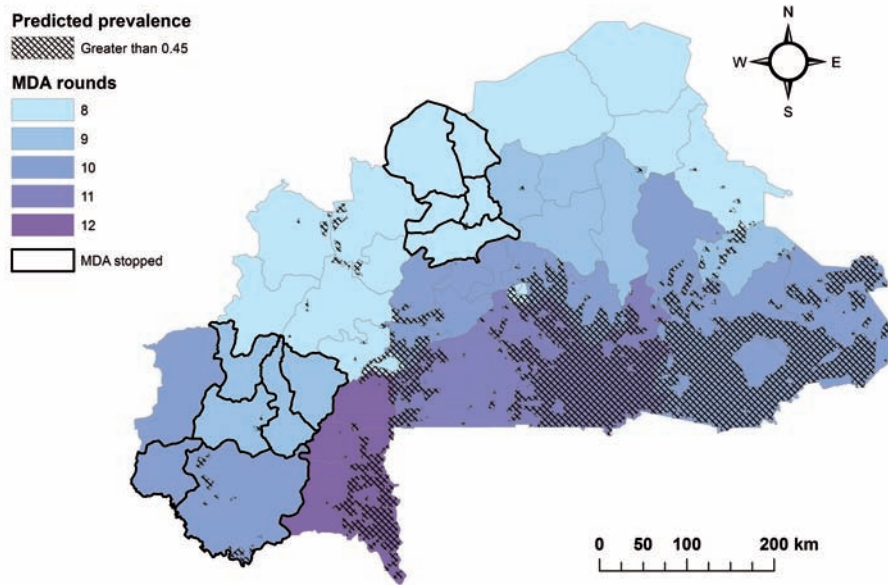Fig. 8. Predictions of prevalence resulting from the fitted model.

Fig. 9. Number of annual rounds of MDA received, up to and including 2012. Outlined areas represent those where MDA has stopped, following successfully reducing LF prevalence to <1%. The shaded area represents a predicted baseline LF prevalence of greater than 45%.

general, elevation in Burkina Faso is low; hence this should not be a limiting factor with regard to the survival of the LF vector. However, low levels of elevation tend to indicate the presence of rivers and lakes, and elevation could therefore be a proxy for this in this model. The presence of cotton crops appears to inhibit LF transmission in this model, which we discuss in more detail below. Further, there is a positive relation- ship between LF prevalence and the number of wet days, SD of NDVI and population density. A long wet season may result in a longer LF transmission period and the variability in vegetation could indicate that the area is neither persistently dry or hot nor persistently wet and densely vegetated. Finally, the higher the num- ber of people in an area, the greater the risk of trans- mission for any disease.
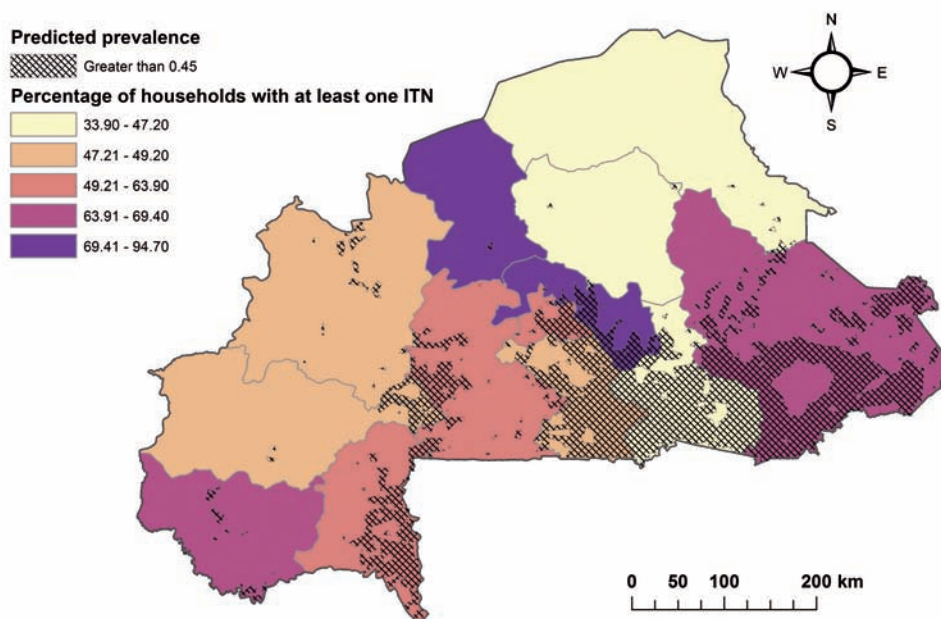

Fig. 10. Proportion of households surveyed in 2010 with at least one ITN. The shaded area represents a predicted baseline LF preva- lence of greater than 45%.

These results support previous studies into the relationship between the environment and prevalence conducted in sub-Saharan Africa (Brengues, 1975; Lindsay and Thomas, 2000; Kelly-Hope et al., 2006) and studies in India (Sabesan et al., 2006). The baseline LF risk map highlights areas where the LF risk is high and accentuates areas where current MDA activities should be sufficient in eliminating the disease as a public health problem. In health districts where the risk fluctuates, it is possible that poorly located prevalence survey sites may result in LF risk being underestimated. This map could therefore be used to guide the targeting of future survey sites to ensure that LF no longer persists at the sub-district level.

It was noted that ITN coverage during the study baseline period was relatively low and also very variable geographically although no suitable contemporaneous data were available. Results from the DHS survey conducted in 2003 showed that the percentage of households owning at least one of any type of net was 40.4%, a figure that was reduced to 4.6% when at least one of these was of the ITN type. However, due to the introduction of a national strategic plan in 2001, these figures are unlikely to represent the situation prior to this time point (Institut National de la Statistique et de la Démographie, 2004). The most recently conducted DHS survey shows that the number of households owning at least one ITN has increased to 56.9% of, with a range from 33.9% in the Central-North region to 94.7% in the North (Institut National de la Statistique et de la Démographie, 2012). This, coupled with MDA coverage of greater than 70% across all endemic health districts should have a significant impact on LF prevalence. However, large areas with a high baseline ICT prevalence in the South have still not reduced LF prevalence to a sufficiently low level despite having received over twice the minimum number of rounds of ivermectin and albendazole as recommended by WHO, i.e. 12 rounds as opposed to five rounds (WHO, 2012b). As these areas coincide with those that have some of the lowest levels of ITN coverage within the country, this suggests more concentrated efforts will be needed to eliminate the disease, including greater use of effective vector control strategies, increasing the frequency of MDA to twice-yearly (Dembele et al., 2010) or using alternative drugs such as doxycycline (Hoerauf et al., 2003; Taylor et al., 2010). Following recommendations by the Global Alliance to Eliminate Lymphatic Filariasis (GAELF), a twice-yearly dosing regime has already been adopted in four health districts in the South-west, following

persistently high LF prevalence and microfilaria density there (Helen Keller International, 2012).

It was observed that LF prevalence was low in a large area in the West compared to the central and eastern parts, although all three areas are climatically and environmentally similar with respect to temperature, rainfall and vegetation. On exploring this further, it was noted that the western region comprises the primary areas for Burkina Faso's cotton crops, a crop that requires large quantities of insecticides. Indeed, 90% of insecticide use in West Africa is attributed to cotton cultivation (Yadouleton et al., 2011) and this could well have contributed to reduced transmission of local *Anopheles* vectors. The lower levels of prevalence in 2000 could be explained by this coupled with the use of ivermectin close to the Dienkoa River (Borsboom et al., 2003). The intensive use of insecticide has, however, would contribute to the development of insecticide resistance in these cotton growing areas, and may have already done so (Diabate et al., 2002; Dabiré et al., 2009; Namountougou et al., 2012). LF risk in these areas may increase over time and should be therefore be monitored.

Insecticide resistance is widespread in Burkina Faso and may also impact the effectiveness of ITNs (Badolo et al., 2012) though the effect of this on LF transmission is currently unknown. However, it has been suggested that insecticide resistance may inhibit the development of filarial worms (McCarroll and Hemingway, 2002). Further, in considering the chromosomal form of the *An. gambiae* s.s. vector in Burkina Faso and its geographical distribution across the country (Bayoh et al., 2001), there are indications that the *An. gambiae* M form is potentially the more efficient vector than the S form (de Souza et al., 2012). Due to this region being very wet (annual rainfall of 800-1000 mm) and at a slightly higher elevation (300-500 m above the mean sea level) compared to the rest of the country, it is likely to be favoured by the S form, which may also contribute to the low prevalence of LF in this area (Bayoh et al., 2001; Kelly-Hope et al., 2006; Namountougou et al., 2012).

We recognise that the variables used to produce the LF baseline risk map do not fully capture the variability observed in LF prevalence across Burkina Faso. The inclusion of the proxy wealth measure extracted from the 1998-1999 DHS data in the socio-demographic model, namely the proportion of households in the nearest DHS surveyed village in the lowest two wealth quintiles indicated that village-level wealth could explain some of this additional variability. Neither additional DHS-derived measures including

the level of education and access to clean drinking water and sanitation facilities, nor similar census-derived province level measures, were significantly related to LF prevalence. However, these measures need to be further developed and refined in order to draw more rigorous conclusions regarding their relationship with LF, especially as the relationship between poverty and LF and NTDs in general is becoming more of an international priority.

Finally, it is important to highlight that the risk map presented in the analysis is based on data on meteorology, land use and population density measured or estimated for Burkina Faso circa 2000. This map will need to be updated as and when climatic and demographic changes occur over time. For example, as the climate becomes hotter and drier (United States Geological Surveys, 2012), or as populations grow and migrate, it is possible that the geographical distribution of LF within Burkina Faso may be affected and potentially increase in areas that were previously less affected and *vice versa* in other areas. This study provides a foundation for a better understanding of the range of factors that influence LF prevalence, and could therefore assist national programmes to be prepared and able to better react to changes in the LF landscape.

References

Badolo A, Traore A, Jones CM, Sanou A, Flood L, Guelbeogo WM, Sagnon N, 2012. Three years of insecticide resistance monitoring in *Anopheles gambiae* in Burkina Faso: resistance on the rise? Malar J 11, 232.

Bayoh MN, Thomas CJ, Lindsay SW, 2001. Mapping distributions of chromosomal forms of *Anopheles gambiae* in West Africa using climate data. Med Vet Entomol 15, 267-274.

Boatin B, 2008. The Onchocerciasis Control Programme in West Africa (OCP). Ann Trop Med Parasitol 102, 13-17.

Borsboom GJ, Boatin BA, Nagelkerke NJ, Agoua H, Akpoboua KL, Alley EWS, Habbema JDF, 2003. Impact of ivermectin on onchocerciasis transmission: assessing the empirical evidence that repeated ivermectin mass treatments may lead to elimination/eradication in West-Africa. Filaria J 2, 8.

Brengues J, 1975. La filariose de ancroft en Afrique de l'Ouest. ORSTOM No. 79.

Brooker S, 2007. Spatial epidemiology of human schistosomiasis in Africa: risk models, transmission dynamics and control. Trans R Soc Trop Med Hyg, 101, 1-8.

Cressie N, 1993. Estimation of the variogram. In: Statistics for spatial data. Wiley Series in Probability and Mathematical Statistics 69-82 pp.

Dabiré KR, Diabaté A, Namountougou M, Toé KH, Ouari A,

Kengne P, Baldet T, 2009. Distribution of pyrethroid and DDT resistance and the L1014F kdr mutation in *Anopheles gambiae* s.l. from Burkina Faso (West Africa). Trans R Soc Trop Med Hyg 103, 1113-1120.

de Souza DK, Koudou B, Kelly-Hope LA, Wilson MD, Bockarie MJ, Boakye DA, 2012. Diversity and transmission competence in lymphatic filariasis vectors in West Africa, and the implications for accelerated elimination of *Anopheles*-transmitted filariasis. Parasit Vectors 5, 259.

Dembele B, Coulibaly YI, Dolo H, Konate S, Coulibaly SY, Sanogo D, Klion AD, 2010. Use of high-dose, twice-yearly albendazole and ivermectin to suppress *Wuchereria bancrofti* microfilarial levels. Clin Infect Dis 51, 1229-1235.

Diabate A, Baldet T, Chandre F, Akoobeto M, Guiguemde TR, Darriet F, Hougard JM, 2002. The role of agricultural use of insecticides in resistance to pyrethroids in *Anopheles gambiae* s.l. in Burkina Faso. Am J Trop Med Hyg 67, 617-622.

Elnaiem DEA, Schorscher J, Bendall A, Obsomer V, Osman ME, Mekkawi AM, Thomson MC, 2003. Risk mapping of visceral leishmaniasis: the role of local variation in rainfall and altitude on the presence and incidence of kala-azar in eastern Sudan. Am J Trop Med Hyg 68, 10-17.

Erlanger TE, Keiser J, Castro MC, Bos R, Singer BH, Tanner M, Utzinger J, 2005. Effect of water resource development and management on lymphatic filariasis, and estimates of populations at risk. Am J Trop Med Hyg 73, 523-533.

Gazzinelli A, Correa-Oliveira R, Yang GJ, Boatin BA, Kloos H, 2012. A research agenda for helminth diseases of humans: social ecology, environmental determinants, and health systems. PLoS Negl Trop Dis 6, e1603.

Gyapong JO, Gyapong M, Evans DB, Aikins MK, Adjei S, 1996. The economic burden of lymphatic filariasis in northern Ghana. Ann Trop Med Parasitol 90, 39-48.

Gyapong JO, Kyelem D, Kleinschmidt I, Agbo K, Ahouandogbo F, Gaba J, Remme J, 2002. The use of spatial analysis in mapping the distribution of bancroftian filariasis in four West African countries. Ann Trop Med Parasitol 96, 695-705.

Gyapong J, Remme J, 2001. The use of grid sampling methodology for rapid assessment of the distribution of bancroftian filariasis. Trans R Soc Trop Med Hyg 681-686.

Hastie T, Tibshirani R, 1986. Generalized additive models. Stat Sci 1, 297-310.

Helen Keller International, 2012. Operational plan 2013 for neglected tropical diseases control in Burkina Faso.

Hoerauf A, Mand S, Fischer K, Kruppa T, Marfo-Debrekyei Y, Debrah AY, Büttner DW, 2003. Doxycycline as a novel strategy against bancroftian filariasis-depletion of *Wolbachia* endosymbionts from *Wuchereria bancrofti* and stop of microfilaria production. Med Microbiol Immun 192, 211-216.

Hosmer DW, Lemesbow S, 1980. Goodness of fit tests for the multiple logistic regression model. Commun Stat - Theor M 9, 1043-1069.

Institut National de la Statistique et de la Démographie, 1996. Recensement general de la population et de l'habitation de 1996 du Burkina Faso. Available at: http://www.insd.bf/fr/spip.php?article310 (accessed on February 2013).

Institut National de la Statistique et de la Démographie, 2000. Burkina Faso Enquête Démographique et de Santé 1998-1999. Available at: http://www.measuredhs.com/pubs/pdf/FR110/FR110.pdf (accessed on February 2013).

Institut National de la Statistique et de la Démographie, 2004. Burkina Faso Enquête Démographique et de Santé 2003. Available at: http://www.measuredhs.com/pubs/pdf/FR154/FR154.pdf (accessed on April 2013).

Institut National de la Statistique et de la Démographie, 2012. Burkina Faso Enquête Démographiqe et de Santé et à Indicateurs Multiples. Available at: http://www.measuredhs.com/publications/publication-FR256-DHS-Final-Reports.cfm (accessed on April 2013).

Kelly-Hope LA, Diggle PJ, Rowlingson BS, Gyapong JO, Kyelem D, Coleman M, Molyneux DH, 2006. Negative spatial association between lymphatic filariasis and malaria in West Africa. Trop Med Int Health 112, 129-135.

Kyelem D, 2007. Epidemiology and control of lymphatic filariasis in Burkina Faso. PhD thesis, Liverpool School of Tropical Medicine.

Kyelem D, Medlock J, Sanou S, Bonkoungou M, Boatin B, Molyneux DH, 2005. Impact of long-term (14 years) bi-annual ivermectin treatment on *Wuchereria bancrofti* microfilaraemia. Trop Med Int Health 10, 1002-1004.

Kyelem D, Sanou S, Boatin B, Medlock J, Coulibaly S, Molyneux DH, 2003. Impact of long-term ivermectin (Mectizan) on *Wuchereria bancrofti* and *Mansonella perstans* infections in Burkina Faso: strategic and policy implications. Ann Trop Med Parasitol 97, 827-838.

Lindsay S, Thomas C, 2000. Mapping and estimating the population at risk from lymphatic filariasis in Africa. Trans R Soc Trop Med Hyg 94, 37-45.

Machault V, Vignolles C, Borchi F, Vounatsou P, Pages F, Briolant S, Rogier C, 2011. The use of remotely sensed environmental data in the study of malaria. Geospat Health 5, 151-168.

McCarroll L, Hemingway J, 2002. Can insecticide resistance status affect parasite transmission in mosquitoes? Insect Biochem Molec 32, 1345-1351.

Molesworth AM, Cuevas LE, Connor SJ, Morse AP, Thomson MC, 2003. Environmental risk and meningitis epidemics in Africa. Emerg Infect Dis, 9, 1287-1293.

Molyneux DH, 2008. Combating the "other diseases" of MDG 6: changing the paradigm to achieve equity and poverty reduction? Trans R Soc Trop Med Hyg 102, 509-519.

Mwobobia IK, Mitsui Y, 1999. Demographic and socio-economic factors with implications for the control of lymphatic filariasis in Kwale District, Kenya. E Afr Med J 76, 495-498.

Namountougou M, Simard F, Baldet T, Diabaté A, Ouédraogo JB, Martin T, Dabiré RK, 2012. Multiple insecticide resistance in *Anopheles gambiae* s.l. populations from Burkina Faso, West Africa. PLoS One 7, e48412.

Odiit M, Bessell PR, Fèvre EM, Robinson T, Kinoti J, Coleman PG, Woolhouse MEJ, 2006. Using remote sensing and geographic information systems to identify villages at high risk for rhodesiense sleeping sickness in Uganda. Trans R Soc Trop Med Hyg 100, 354-362.

Programme National d'Elimination de la Filariose Lymphatique Burkina Faso, 2012. Rapport Annuel.

Sabesan S, Raju HKK, Srividya A, Das PK, 2006. Delimitation of lymphatic filariasis transmission risk areas: a geo-environmental approach. Filaria J 5, 12.

Taylor MJ, Hoerauf A, Bockarie M, 2010. Lymphatic filariasis and onchocerciasis. Lancet 376, 1175-1185.

The Global Atlas of Helminth Infections, 2012. Available at: http://www.thiswormyworld.org/maps/burkina-faso (accessed on February 2013).

United States Geological Surveys, 2012. A climate trend analysis of Burkina Faso. Available at: http://pubs.usgs.gov/fs/2012/3084/fs2012-3084.pdf (accessed on April 2013).

Upadhyayula SM, Mutheneni SR, Kadiri MR, Kumaraswamy S, Nagalla B, 2012. A cohort study of lymphatic filariasis on socio economic conditions in Andhra Pradesh, India. PLoS One 7, e33779.

Wood S, 2006, Generalised additive models an introduction with R. Chapman and Hall/CRC; Taylor and Francis Group. 121-144 pp.

WHO, 2010. LF progress report and strategic plan 2010-2020 of the Global Programme to Eliminate Lymphatic Filariasis: halfway towards eliminating lymphatic filariasis.

WHO, 2012a. Lymphatic filariasis factsheet no. 102. Available at: http://www.who.int/mediacentre/factsheets/fs102/en/ (accessed on April 2013).

WHO, 2012b. Transmission assessment surveys in the Global Programme to Eliminate Lymphatic Filariasis: WHO position statement. Wky Epidemiol Rec 87, 478-482.

Yadouleton A, Martin T, Padonou G, Chandre F, Asidi A, Djogbenou L, Akogbeto M, 2011. Cotton pest management practices and the selection of pyrethroid resistance in *Anopheles gambiae* population in northern Benin. Parasite Vector 4, 60.

Yang K, Sun LP, Huang YX, Yang GJ, Wu F, Hang DR, Zhou XN, 2012. A real-time platform for monitoring schistosomiasis transmission supported by Google Earth and a web-based geographical information system. Geospat Health 6, 195-203.