Geostatistical Modeling of Malaria Endemicity using Serological Indicators of Exposure Collected through School Surveys

Ruth A. Ashton,* Takele Kefyalew, Alison Rand, Heven Sime, Ashenafi Assefa, Addis Mekasha, Wasihun Edosa, Gezahegn Tesfaye, Jorge Cano, Hiwot Teka, Richard Reithinger, Rachel L. Pullan, Chris J. Drakeley, and Simon J. Brooker

Malaria Consortium, London, United Kingdom; Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom; Malaria Consortium, Addis Ababa, Ethiopia; Ethiopian Public Health Institute, Addis Ababa, Ethiopia; Oromia Regional Health Bureau, Addis Ababa, Ethiopia; President's Malaria Initiative, United States Agency for International Development, Addis Ababa, Ethiopia; RTI International, Washington, District of Columbia

Abstract. Ethiopia has a diverse ecology and geography resulting in spatial and temporal variation in malaria transmission. Evidence-based strategies are thus needed to monitor transmission intensity and target interventions. A purposive selection of dried blood spots collected during cross-sectional school-based surveys in Oromia Regional State, Ethiopia, were tested for presence of antibodies against *Plasmodium falciparum* and *P. vivax* antigens. Spatially explicit binomial models of seroprevalence were created for each species using a Bayesian framework, and used to predict seroprevalence at 5 km resolution across Oromia. School seroprevalence showed a wider prevalence range than microscopy for both *P. falciparum* (0–50% versus 0–12.7%) and *P. vivax* (0–53.7% versus 0–4.5%), respectively. The *P. falciparum* model incorporated environmental predictors and spatial random effects, while *P. vivax* seroprevalence first-order trends were not adequately explained by environmental variables, and a spatial smoothing model was developed. This is the first demonstration of serological indicators being used to detect large-scale heterogeneity in malaria transmission using samples from cross-sectional school-based surveys. The findings support the incorporation of serological indicators into periodic large-scale surveillance such as Malaria Indicator Surveys, and with particular utility for low transmission and elimination settings.

INTRODUCTION

With the rekindling of malaria elimination goals,[1,2] there is an increased need to quantify patterns of and changes in malaria risk to support evidence-based targeting of interventions, implement surveillance strategies to monitor changes in transmission intensity, and assess feasibility of local elimination.[3,4]

Low-transmission settings present specific challenges to implementation of cross-sectional surveys: 1) highly seasonal transmission can result in underestimates of population parasite rates if sampling does not occur during the peak transmission period; 2) low-density infections are frequent and either will be underestimated by microscopy or rapid diagnostic tests (RDTs),[5,6] or require significantly increased resources to screen samples by polymerase chain reaction (PCR); and 3) where diagnostic data are used to develop spatial prediction models, there is a risk that the true extent of transmission will be underestimated since recently cleared and low-density infections will not be included. While new strategies such as reactive case detection[7–9] and "rolling" cross-sectional surveys[10] have been trialed, there remains a need to develop strategies to track changes in low and unstable transmission settings.

Detection of anti-*Plasmodium* antibodies eluted from dried blood spots has been demonstrated to be robust,[11,12] yielding estimates of seroprevalence and seroconversion rates that are representative of malaria transmission intensity within a community.[13,14] Because antibodies persist after infection clearance, they offer the opportunity to examine exposure to malaria over a wider period than is typically possible through detection of parasitemia during a cross-sectional survey by means of microscopy, RDTs, or PCR-based methods.

Serological indicators are increasingly being used in community-based malaria epidemiological studies to assess small- and large-scale spatial heterogeneities of and changes in transmission.[13,15–19] Schools provide a useful alternative platform for collection and monitoring of malariometric indicators, offering logistical advantages (e.g., simplified selection of participants, high compliance, and reduced survey costs) over standard community-based cross-sectional surveys.[20–22] Although schools consistently yield higher estimates, school-survey seroprevalence estimates have repeatedly been shown to strongly correlate with community-survey seroprevalence.[23]

This study explored the use of serological indicators collected from a large-scale school-based survey to describe differences in *P. falciparum* and *P. vivax* endemicity in a low-transmission setting. Spatially explicit Bayesian modeling techniques were used to explore relationships between serological indicators at population level and explanatory environmental variables, to predict estimated endemicity levels at subnational scale.

METHODS

Study setting. Ethiopia has a diverse ecology, and malaria transmission is known to be spatially heterogeneous, related to variables such as altitude, temperature, rainfall, and presence of local water bodies or dams.[24–28] Malaria transmission is temporally variable because of seasonal rainfall, with a major transmission season from September to December and a minor transmission season from April to May. Cases are due to both *P. falciparum* and *P. vivax*. The Malaria Indicator Survey in 2011 demonstrated a low parasite prevalence within the population living in malaria-risk areas, estimated at 1.3% by microscopy and 4.5% by RDT in areas < 2,000 m.[29]

Survey data. Data presented in this paper are drawn from a large cross-sectional survey conducted in 197 government primary schools in Oromia Regional State, Ethiopia, in 2009.[30] Full detail of school and child selection as well as sample collection has been presented elsewhere.[30] In brief, at each school 55 girls and 55 boys were randomly selected.

* Address correspondence to Ruth A. Ashton, Malaria Consortium, Development House, 56-64 Leonard Street, London. E-mail: r.ashton@malariaconsortium.org

TABLE 1
Number of schools and children tested by ELISA against each antigen, stratified by school selection criteria: high microscopy prevalence, high anemia prevalence, and randomly selected

	Any *P. falciparum* antigen				Any *P. vivax* antigen			
	Schools		Children		Schools		Children	
Total tested	62		5,913		71		6,609	
	*Pf*MSP-1		*Pf*GLURP		*Pv*MSP-1		*Pv*AMA	
School selection criteria	Schools	Children	Schools	Children	Schools	Children	Schools	Children
High microscopy prevalence	20	2,088	20	2,093	20	2,080	20	2,074
High anemia prevalence	20	2,118	20	2,092	20	2,080	20	2,104
Random selection	22	1,614	10	1,037	31	2,327	10	1,024
Total tested	62	5,820	50	5,222	71	6,487	50	5,202

AMA = apical membrane antigen; ELISA = enzyme-linked immunosorbent assay; GLURP = glutamate-rich protein; MSP = merozoite surface protein; *Pf* = *Plasmodium falciparum*; *Pv* = *P. vivax*.

They provided finger-prick blood samples for preparation of thick- and thin-blood smears, hemoglobin measurement (HemoCue Ltd., Angelhölm, Sweden), and collection of blood spots on filter paper (Whatman 3MM; Whatman, Maidstone, United Kingdom). School location was measured using a global positioning satellite receiver (eTREX; Garmin International, Olathe, KS).

For serological analysis samples were selected purposively from 1) 20 schools with highest prevalence of *Plasmodium* infection detected by microscopy (range 0.9–14.5%); 2) 20 schools with highest proportion of anemic (classified according to WHO,[31] including adjustment by altitude) children (range 34.2–51.4%); and 3) a random selection of remaining schools surveyed (Table 1). Purposive selection was conducted to capture a range of transmission settings, and since resources were not available to complete enzyme-linked immunosorbent assay (ELISA) on all blood spots collected during surveys.

Enzyme-linked immunosorbent assay. Blood spots from 50 schools were analyzed in London, United Kingdom, against *P. falciparum* merozoite surface protein-1$_{19}$ (*Pf*MSP-1), *P. falciparum* glutamate-rich protein-R2 (*Pf*GLURP), *P. vivax* MSP 1$_{19}$ (*Pv*MSP-1), and *P. vivax* apical membrane antigen-1 (*Pv*AMA). In Addis Ababa, Ethiopia, blood spots from a further 12 schools were analyzed against *Pf*MSP-1 and 21 schools against *Pv*MSP-1.

Antibodies were eluted from dried blood spots, and samples were tested for IgG against *P. falciparum* and *P. vivax* antigens according to methods described elsewhere.[12] Duplicate optical density (OD) values with > 20% variation were excluded. Raw ODs were corrected by blank OD and normalized between plates by fitting to the midpoint of a standard curve produced by serial dilution of hyperimmune serum (i.e., pooled hyperimmune serum from Tanzania for *P. falciparum*, and reconstituted *P. vivax* and *P. malariae* hyperimmune serum (NIBSC code 72/096, Hertforshire, UK) for *P. vivax*). Normalized ODs and identification numbers were exported into Stata 12.0 (STATA Corporation, College Station, TX). Individual samples were classified as seropositive or seronegative against each antigen using a mixture model, whereby the mean of the seronegative distribution plus three standard deviations was defined as the seropositive cutoff.[12,13] Binary variables were generated to describe, in summary, seropositivity by species: for example, *P. falciparum* seropositive samples were defined as seropositive against either *Pf*MSP-1 and/or *Pf*GLURP. In the absence of a gold standard for anti-*Plasmodium* antibody detection, it is not possible to determine the sensitivity or specificity of the ELISA, but the mixture model approach is commonly used in low-transmission settings,[15,32–34] where the population is expected to include true seronegative and true seropositive individuals. Various nonlinear regression functions were fitted to scatter plots of microscopy prevalence and seroprevalence for each species separately using least squares regression, to describe the relationship between microscopy and serology.

Remote sensing environmental data. Elevation data were extracted from the Shuttle Radar Topography Mission (SRTM) digital elevation model at 90 m resolution,[35] resampled to 250 m, and further processed to estimate slope in degrees. Gridded precipitation and temperature data at 1 km resolution were extracted from preprocessed data available on WorldClim.[36,37] This source provides a set of data layers generated through interpolation of average monthly climate data obtained during the period 1950–2000. Euclidean distance to water bodies was calculated using SRTM Water Body Data files at 250 m resolution,[38] and distance to rivers and roads calculated using data from Digital Chart of the World at 250 m resolution.[39] Land cover type was extracted from the qualitative global land cover map for 2005, defined within the United Nations land cover classification system using environmental satellite (ENVISAT) mission's Medium Resolution Imaging Spectrometer (MERIS) sensor at 300 m resolution.[40] Monthly raster datasets of normalized difference vegetation index (NDVI) indicators at 1 km resolution were extracted from the SPOT 5 vegetation project[41] for the period 2005 and 2009. Gridded mean, minimum, maximum, and standard deviations were obtained for the entire period and for the specific survey year (2009). Population density was extracted from the AfriPop project at 100 m resolution,[42] and rural–urban classification at 1 km from the Global Rural-Urban Mapping Project (GRUMP).[43] Input grids were either extended or clipped to match the geographic extent of a land mask template, aligned to it, and eventually resampled to a 5 km resolution by bilinear interpolation for continuous surface, and majority approach for categorical data. Bilinear interpolation determines the new value of a cell based on a weighted distance average of the nearest input cell centers, whereas majority approach determines the new value based on the most popular values within the resampling window. We assumed that both methods cause some smoothing of the data. Environmental data were extracted to school locations using ArcMap 12.0 (Environmental Systems Research Institute Inc., Redlands, CA).

Model development and testing. Environmental and serology data were merged and analyzed using Stata 12.0. Continuous environmental variables were standardized to facilitate later model convergence. Models were developed separately to describe *P. falciparum* or *P. vivax* seroprevalence.

Univariate associations between school seroprevalence and environmental variables were explored, and colinearity (correlation coefficient > 0.9) between variables was tested. A school-level minimal adequate logistic regression model was developed by the backward stepwise method, whereby variables with $P > 0.05$ were removed in the order of least significance; all excluded variables were subsequently retested in the final model. Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used to inform model selection.[44,45]

Four multivariate Bayesian binomial regression models were developed using WinBUGS 1.4 (Medical Research Council Biostatistics Unit and Imperial College London, London, United Kingdom) for *P. falciparum* and for *P. vivax.* The most complex model included the retained school-level environmental variables, school-level random effect, and school-level geostatistical random effect (using an isotropic, stationary exponential decay function).[46] Additional models excluded the environmental variables, the spatial random effect, or both.

Semi-informative priors were set for the rate of decay of spatial correlation, ϕ, informed by the maximum and minimum distance between schools, and non-informative priors used for other coefficients. Models were burned in for 10,000 iterations to achieve convergence, and then nodes were sampled for 10,000 iterations, thinning each 10 iterations. Final model selection was informed by examining the variance of school and spatial random effects and deviance information criterion (DIC).[47]

Model validation. Models were externally validated by training the model on an N-5 school dataset, then predicting probability that seroprevalence thresholds (2%, 5%, and 40%) exceeded for the five excluded schools. These thresholds were estimated to describe the lowest and highest areas of seroprevalence, and hence endemicity, to support specific intervention targeting. The process was repeated until predictions for all schools were available. Model performance was assessed by examining the area under the curve (AUC) of the receiver operator characteristic (ROC) at each threshold.[48] AUC > 0.7 indicates a reasonable discriminative capacity, and AUC > 0.9 very good discriminative capacity.[49,50]

Generating a predictive seroprevalence map. A grid of 12,048 locations at 5 km spacing was generated across Oromia, and environmental variables included in final models were extracted to these locations. The selected Bayesian models were trained on actual school seroprevalence data, then predicted at each location by calculating the sum of the products of the covariate coefficients and the values of the covariates at each grid node, plus the interpolated geostatistical random effect, and back transforming from the logit to the prevalence scale.

Ethical considerations. The school surveys received ethical clearance from the Ethiopian Science and Technology Agency (RDHE/2-89/2009), with additional clearance subsequently given for serological analysis of blood spots (3.10/53/2003). Consent for participation used a passive, opt-out procedure, with school director providing written consent for the survey to proceed. Schools were requested to hold meetings in advance with parents to inform them of the survey and allow withdrawal of children if necessary. Participating children gave written assent and were informed of their right to withdraw at any time. Children reporting fever during surveys were tested with a multispecies HRP2-panLDH RDT (CareStart; AccessBio, Somerset, NJ), and any child with a positive RDT was treated according to the national guidelines.[51]

RESULTS

Serology findings. Serology results were available for *P. falciparum* from 5,914 children from 62 schools, with a mean 95 (range 10–111) samples per school. *Plasmodium vivax* results were available from 6,609 children from 71 schools, with mean 93 (range 5–111) samples per school. Data were from children aged 5 to 18 years (mean 11 years).

Of all children tested, 11.6% (688/5,913) were *P. falciparum* seropositive and 11.1% (735/6,609) *P. vivax* seropositive; 1.0% and 0.5% children were microscopy-positive for *P. falciparum* and *P. vivax* parasites, respectively. Cross-tabulation of microscopy and antigen-specific serology results are presented in Table 2. Where data were available for both species, 4.7% of 5,420 children were seropositive against both species. When restricting our analyses to schools with more than 50 children tested (56 schools for *P. falciparum*, 62 for *P. vivax*), *P. falciparum* and *P. vivax* school seroprevalence ranged from 0% to 50% and 0% to 53.7%, respectively.

Among 50 schools tested against 4 antigens, correlation was seen between school seroprevalence determined for *Pf*MSP-1 and *Pf*GLURP ($R^2 = 0.84$), and for *Pv*MSP-1 and *Pv*AMA ($R^2 = 0.80$). For both species, coating plates with MSP-1 resulted in higher sensitivity than *Pf*GLURP or *Pv*AMA. A strong correlation ($R^2 = 0.84$) was seen between school *P. falciparum* and *P. vivax* seroprevalence.

Comparing serology to microscopy. Schools with 0% positive samples by microscopy were found to have from 0% to 30% seroprevalence. Although the proportion of microscopy positive and seropositive children in a school are not directly comparable, it is plausible to expect some association between the two measures (Figure 1).

TABLE 2

Description of frequency of diagnostic test (microscopy and serology) results at individual level (combinations of microscopy and seropositivity by antigen are presented for *Plasmodium falciparum* and *P. vivax* separately)

P. falciparum diagnostic tool combinations			
		*Pf*GLURP +	*Pf*GLURP −
Microscopy *Pf* +	*Pf*MSP-1 +	38	5
Microscopy *Pf* +	*Pf*MSP-1 −	3	6
Microscopy *Pf* −	*Pf*MSP-1 +	217	246
Microscopy *Pf* −	*Pf*MSP-1 −	106	4,481
P. vivax diagnostic tool combinations			
		*Pv*AMA +	*Pv*AMA −
Microscopy *Pv* +	*Pv*MSP-1 +	4	10
Microscopy *Pv* +	*Pv*MSP-1 −	2	14
Microscopy *Pv* −	*Pv*MSP-1 +	141	381
Microscopy *Pv* −	*Pv*MSP-1 −	78	4,423

AMA = apical membrane antigen; GLURP = glutamate-rich protein; MSP = merozoite surface protein; *Pf* = *P. falciparum*; *Pv* = *P. vivax*.
Data are only presented for individuals with results recorded for *P. falciparum* microscopy, *Pf*GLURP and *Pf*MSP-1 ($N = 5,102$), and individuals with complete results for *P. vivax* microscopy, *Pv*AMA and *Pv*MSP-1 ($N = 5,053$).

Environmental risk factors. Colinearity was found among the precipitation, temperature and NDVI variables, and between distance to both permanent and any type of water body. Distance to water bodies and rivers, land gradient, and urban areas showed univariate associations with both *P. falciparum* and *P. vivax* seroprevalence (Supplemental Table 1).

The minimal adequate multivariate *P. falciparum* frequentist model includes elevation and angle of land slope, distance to permanent river, bare or sparse land cover, population density, and urban areas. The minimal adequate *P. vivax*

A

B
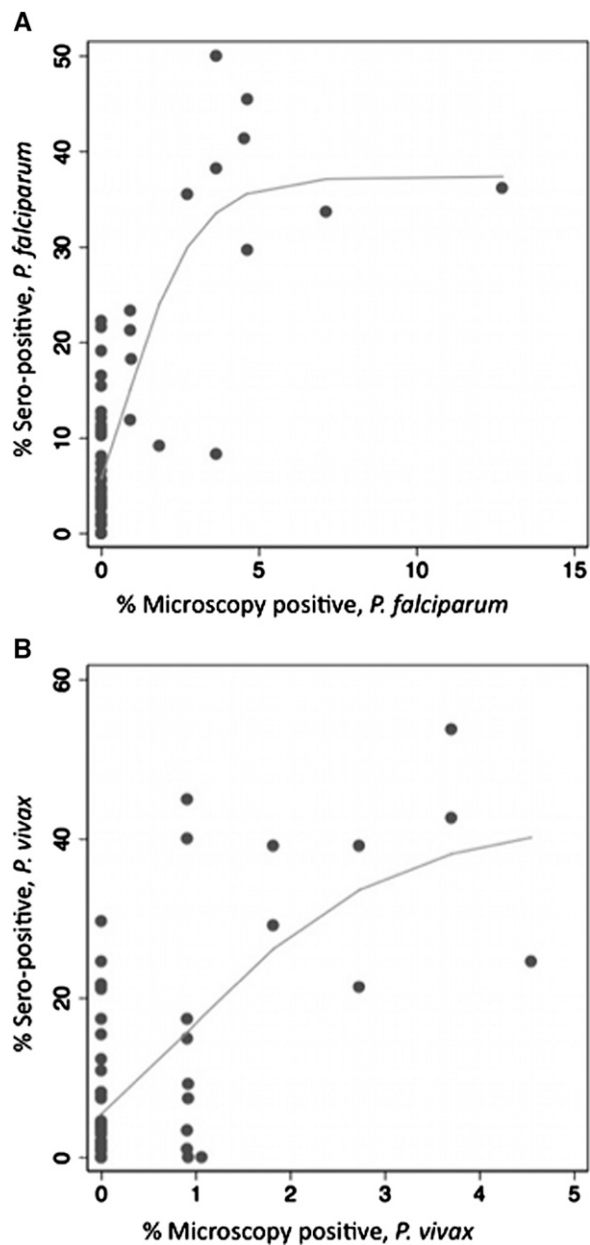
FIGURE 1. School-level seroprevalence and prevalence of infection detected by microscopy for *Plasmodium falciparum* (**A**) and *P. vivax* (**B**). Scatter plots are presented for 56 schools with *P. falciparum* data and 62 schools with *P. vivax* data, restricted to those with serology results from ≥ 50 children. Nonlinear regression identified a Gompertz function as best fit to *P. falciparum* ($R^2 = 0.810$), and to *P. vivax* data ($R^2 = 0.657$).

frequentist model includes distance to permanent river and water body, precipitation during the wettest quarter of the year (projection from 1950 to 2000 to allow for high spatial resolution), and mean NDVI over the preceding 5 years.

Bayesian modeling of *P. falciparum*. When comparing output from nonspatial and spatial models of *P. falciparum*, incorporating spatial structure in models was found to explain much of the variation between schools, indicated by a reduction of σ^2_{school} when spatial random effects were included in models. A lower DIC in models including spatial structure justified retention in the final *P. falciparum* Bayesian model.

Inclusion of environmental variables in the spatial model was shown to reduce $\sigma^2_{spatial}$ and the DIC, as well as increase the rate of decay of spatial correlation (φ), indicating that much of the first-order spatial variation can be explained adequately by the included environmental data. Therefore, the final model to describe *P. falciparum* seroprevalence in Oromia incorporates environmental covariates to explain first-order deterministic spatial variation, with the spatial random effect adequately capturing second-order structure (Table 3). The final model had mean prediction error of −0.31, indicating a tendency to moderately underpredict *P. falciparum* seroprevalence. Mean absolute error of the model, indicating the magnitude of error in predictions, was 6.64. Internal validation demonstrates a good discriminatory ability of the final model for 2% and 5% seroprevalence thresholds, with an AUC of 0.83 and 0.84, respectively. The model performs very well in identifying areas of over 40% seroprevalence (AUC = 0.96). Actual and predicted school seroprevalence were found to be correlated (Pearson $r = 0.62$, $P < 0.001$). The final model was used to predict *P. falciparum* seroprevalence at 5 km resolution across Oromia, the posterior mean is shown in Figure 2, with probability of 2%, 5%, and 40% thresholds shown in Figure 3.

Bayesian modeling of *P. vivax*. Similar to the *P. falciparum* models, incorporating spatial structure in *P. vivax* models was found to explain much of the variation between schools, indicated by a reduction of σ^2_{school} and lower DIC. Inclusion

TABLE 3
Final Bayesian *Plasmodium falciparum* model developed using data from 62 schools, and *P. vivax* model developed from 71 schools' data (both models retained school-level and spatial random effects)

	P. falciparum model parameter value (95% BIC)	*P. vivax* model parameter value (95% BCI)
Altitude	−0.568 (−1.035, −0.087)	–
Slope	−0.595 (−0.996, −0.234)	–
Distance to permanent river	−0.411 (−0.774, −0.036)	–
Population density in 2010	0.418 (−0.107, 0.911)	–
Bare or sparse land (binary)	1.026 (−0.392, 2.298)	–
Urban area (binary)	−3.13 (−6.279, −0.028)	–
σ^2_{school}	0.254 (0.006, 1.250)	0.288 (0.013, 0.939)
$\sigma^2_{spatial}$	1.183 (0.177, 2.311)	3.631 (1.31, 10.85)
ϕ	9.763 (2.631, 19.03)	0.866 (0.211, 2.093)
Range in kilometer	45.57 (17.54, 127)	548.3 (160.5, 1,592)
DIC	308.8	330.8

BIC = Bayesian information criterion; DIC = deviance information criterion.
School and spatial variance (σ^2_{school} and $\sigma^2_{spatial}$), rate of decay of spatial correlation (φ), range in kilometer at which correlation between schools falls to 5% are presented with 95% Bayesian credible intervals. The *P. falciparum* model includes parameter values and 95% BCI for standardized environmental fixed effects. No environmental fixed effects were retained in the final *P. vivax* model.
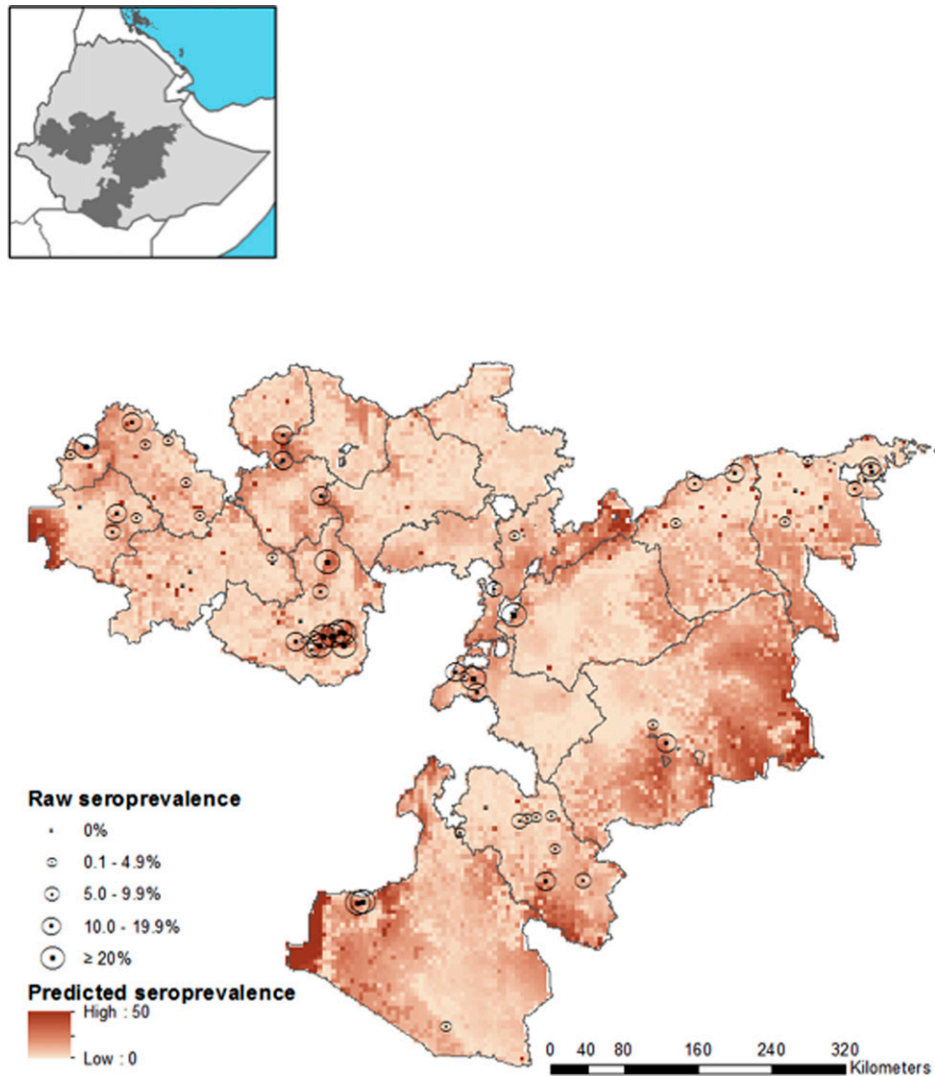
FIGURE 2.   Map of predictive *Plasmodium falciparum* seropositivity using spatial model with environmental fixed effects. Measured *P. falciparum* seroprevalence from the 62 schools used to train the model are shown by circles with size proportional to seroprevalence. Inset map indicates the location of Oromia Regional State (shaded) within Ethiopia.

of environmental variables in the spatial *P. vivax* model did not substantially reduce the $\sigma^2_{spatial}$, and little difference was seen in φ and DIC between the models with and without environmental variables. The final model for *P. vivax* is, therefore, the spatial model with no environmental covariates (Table 3). The final model had mean prediction error of 0.03, and mean absolute error, indicating the magnitude of error in predictions, of 6.64. Internal validation of this model indicates good performance at the 2% seroprevalence threshold (AUC = 0.81) and very good performance at 5% and 40% thresholds (AUC = 0.91 for both). Actual and predicted seroprevalence were correlated (Pearson r = 0.68, P < 0.001). Predictions of the final model at 5 km resolution are displayed in Figure 4 as the posterior mean seroprevalence, and in Figure 5 as the probability of seroprevalence thresholds being exceeded.

DISCUSSION

This study shows the capability of serological markers to detect large-scale heterogeneity in malaria transmission using samples collected during cross-sectional school surveys, in a setting with seasonal and low transmission. Seroprevalence was found to be associated with environmental variables; this relationship was used to predict seroprevalence at unsampled locations using Bayesian geostatistical modeling methods incorporating fixed and random effects.

School seroprevalence determined by different antigens showed strong correlation for each species, with MSP showing higher sensitivity for both *P. falciparum* and *P. vivax* than *P. falciparum* GLURP and *P. vivax* AMA, respectively. However, previous studies have shown AMA-1 to have higher immunogenicity than MSP-1.[11] As transmission declines, individual antibody responses become more disparate, and therefore it is recommended that future serological analysis be conducted using multiple antigens or a whole parasite lysate.

The range of seroprevalence found across schools where no children were microscopy positive during the cross-sectional survey demonstrates the value of serological indicators, that is, in differentiating between schools where transmission occurs but the peak transmission period was missed by surveys, and
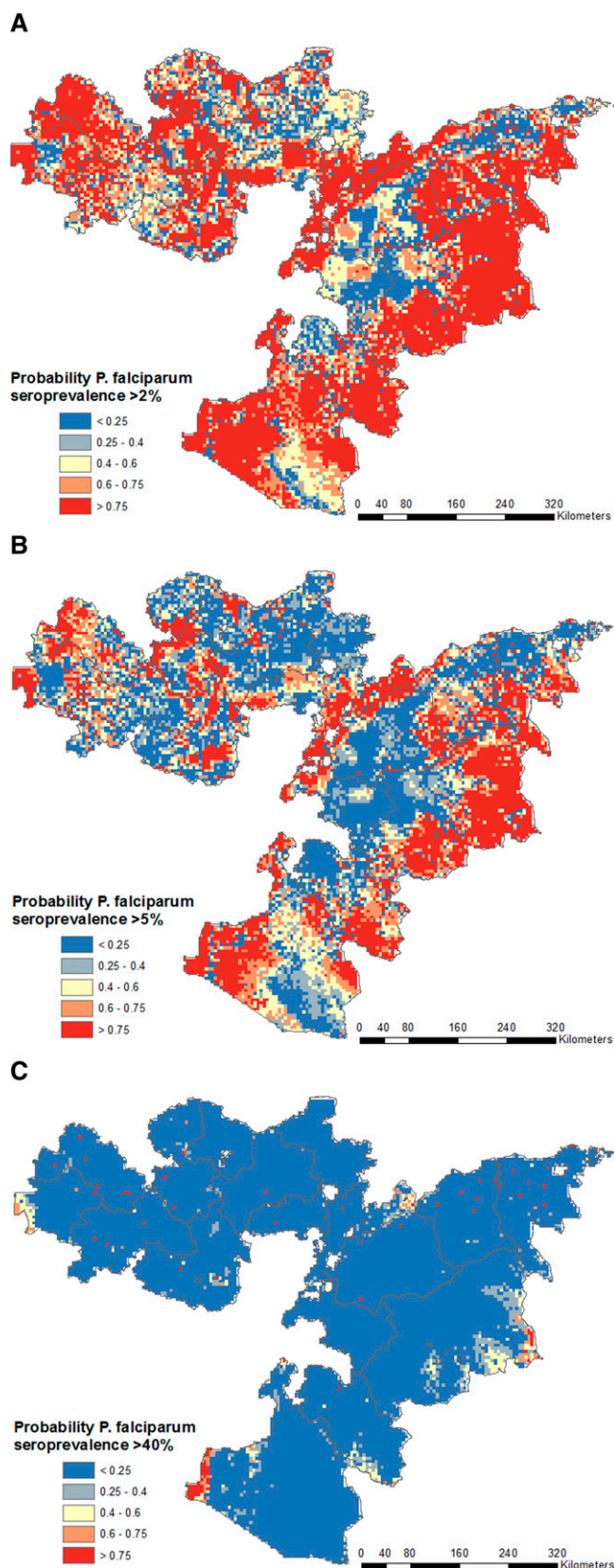
FIGURE 3. Probability of *Plasmodium falciparum* seroprevalence exceeds the defined thresholds of 2% (**A**), 5% (**B**), and 40% (**C**) according to final predictive model for *P. falciparum*. Red areas are those very likely to exceed the threshold, blue areas very unlikely to exceed the threshold, and pale yellow areas have high uncertainty.

those with very low malaria risk. Differences may also be apparent if transmission has ceased in the area in recent years before the age of the youngest school child. This is difficult to demonstrate without clinical data.

Both species' final models incorporated a spatial random effect to describe spatial autocorrelation, whereby schools located closely together were more similar than schools at greater distance. All models included a nonspatial random effect. The *P. falciparum* model indicated that spatial auto-correlation was present to a distance of approximately 46 km, while the *P. vivax* model showed a range of over 500 km. The *P. falciparum* range is a distance at which similarities in climatic factors and ecology would be expected, and therefore it is feasible that these areas experience similar transmission intensity. However, the very large spatial range of *P. vivax* suggests that the spatial random effect is capturing other large-scale variations not tested for inclusion in the Bayesian spatial model. A similar finding was reported in spatial modeling of malaria in Bangladesh, where environmental variables described a large proportion of spatial variation in *P. falciparum*, but little of the *P. vivax* distribution.[52]

Frequentist models developed for *P. vivax* suggested biologically plausible environmental risk factors of distance to rivers and water bodies, vegetation cover, and precipitation; nonetheless, these did not adequately explain the large-scale trends in *P. vivax* seropositivity after accounting for spatial dependency. The final *P. vivax* map presented here therefore simply uses spatial interpolation to predict seroprevalence at unsampled locations. The larger spatial scale of *P. vivax* may be due in part to the production of hypnozoites, since the reactivation of parasites and subsequent antibody production may occur in a different location to site of parasite acquisition, or in the absence of ongoing transmission. However, it is unlikely that recrudescent infections would have had a major confounding effect on school seroprevalence, unless large-scale population movements would have occurred. Furthermore, the wider *P. vivax* range may be due to the parasite's ability to generate sporozoites at lower temperatures and the potential to be transmitted at higher altitudes.[53] Indicators of temperature and altitude, considered to define vector survival and sporogeny, were not found to be associated with *P. vivax* seroprevalence, and not retained in the final multivariate model.

The key environmental variables identified for inclusion in the *P. falciparum* map indicated that higher risk exists in low-altitude and low-gradient areas close to rivers. We postulate that seasonal flooding in flatlands where floodwaters may pool and act as vector breeding sites could be the driver of this relationship.

A further extension to the current models in the future could be incorporation of intervention coverage, such as districts targeted by indoor residual spraying of households with insecticide, and estimations of long-lasting mosquito net coverage and use alongside environmental covariates. Furthermore, a Bayesian approach to selection of environmental covariates may have resulted in a different panel of covariates tested in final models.

Despite the difficulties in modeling *P. vivax* seroprevalence, our maps of both *P. falciparum* and *P. vivax* seroprevalence do show broad concordance with predictive maps developed by the Malaria Atlas Project to describe age-standardized parasite rates using model-based geostatistical prediction methods,[48,54,55] with similar areas of Oromia identified as

Raw seroprevalence
- 0%
- 0.1 - 4.9%
- 5.0 - 9.9%
- 10.0 - 19.9%
- ≥ 20%

Predicted seroprevalence
High : 50
Low : 0
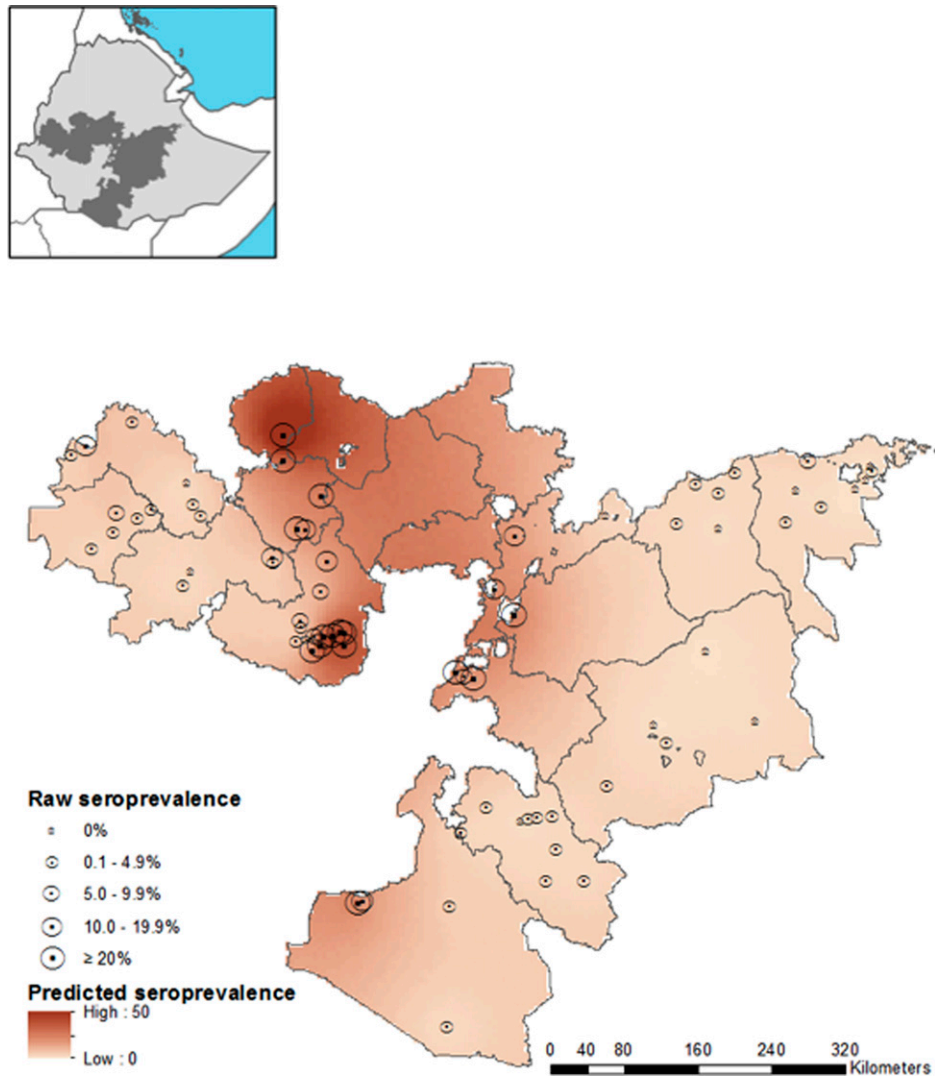
0 40 80   160   240   320
Kilometers

FIGURE 4. Map of predictive *Plasmodium vivax* seropositivity, using spatial model without environmental fixed effects. Measured *P. vivax* seroprevalence from the 71 schools used to train the model are shown by circles with size proportional to seroprevalence. Inset map indicates the location of Oromia Regional State (shaded) within Ethiopia.

areas of highest and lowest risk for malaria. Survey locations with microscopy-positive samples in the most recent Ethiopian Malaria Indicator Survey in 2011 also correlate with our predictive map, with infections identified along the Rift Valley as well as in the far west of Oromia.[29] Considering that sites included in the current modeling approach were selected using a combination of purposive and random methods, a further improvement to the approach presented here would involve defining a likelihood function more appropriate for preferentially sampled geostatistical data, as described in Diggle and others.[56] The use of conventional geostatistical methods to model these data, which assumes non-preferential sampling, could potentially result in misleading inferences. Nevertheless, since a stratified random sampling approach was used for selection of the 200 schools included in the wider malariometric survey, and the use of both purposive and random methods to identify schools for serological testing, these results remain meaningful, even if a cautious interpretation considering potential biases is required.

This study was designed to evaluate large-scale spatial heterogeneity of *P. falciparum* and *P. vivax* malaria. Although logistical constraints limited the number of samples analyzed, the original surveys were powered to microscopy-based parasite rate—therefore, seroprevalence rates being higher than microscopy should mean that adequate samples were examined to evaluate associations with environmental variables and build the statistical model. The study was not designed to assess microheterogeneity in transmission within communities, which has been demonstrated in other settings with similarly low transmission levels (e.g., Somalia, The Gambia, Guinea-Bissau).[15,17] The randomization process and use of school-attending children as a sampling frame should result in a sampled population representative of the whole school catchment area and wider community. We acknowledge that there is potential for school catchment areas in Ethiopia to have diversity in transmission intensity as a result of steep gradients and presence of local water bodies, dams, and irrigation systems.[24,26] Individual differences in immune status
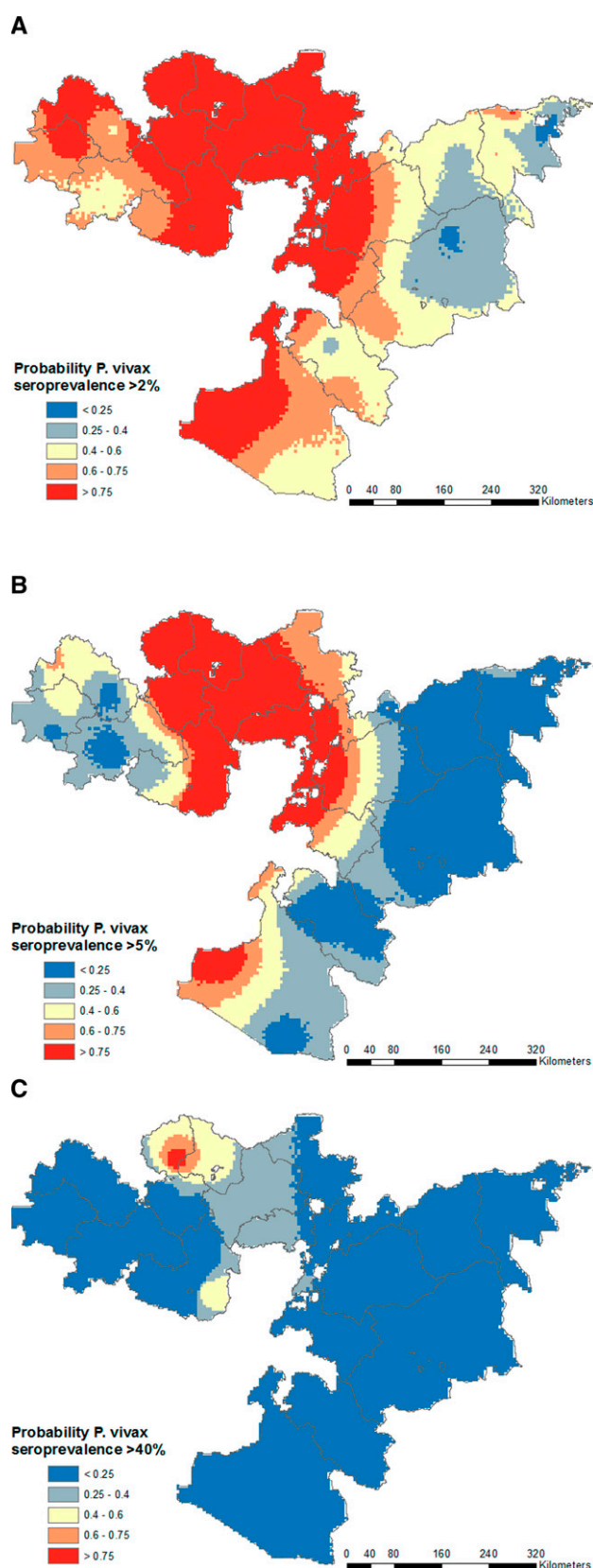
FIGURE 5. Probability of *Plasmodium vivax* seroprevalence exceeds the defined thresholds of 2% (**A**), 5% (**B**), and 40% (**C**) according to final predictive model for *P. vivax*. Red areas are those very likely to exceed the threshold, blue areas very unlikely to exceed the threshold, and pale yellow areas have high uncertainty.

and antibody production in response to *Plasmodium* antigen exposure are expected, and may be moderated by other parasitic infections, including helminthes[57]; yet, infection risks for these are likely to be broadly similar across all sites, and individual differences in immune response are likely to be randomly dispersed among the population.

The Bayesian spatially explicit models developed in this study could be refined by inclusion of serology data from additional sites, both within Oromia to assist in categorizing areas of high model uncertainty, as well as from other regional states to assist in developing a nationally representative risk map. Serological analysis of filter paper blood spots included in periodic national surveys such as Malaria Indicator Surveys or Demographic and Health Surveys would be a simple strategy to collect additional seroprevalence data nationally.

Generation of estimates of cluster seroprevalence can therefore complement the collection of parasitological indicators from periodic large scale surveys in low-transmission settings by indicating recent and historical transmission intensity, depending on population tested and antigens used. In settings with unstable transmission, these data may indicate receptivity to transmission,[58,59] therefore can support policy makers in targeting interventions to areas of current transmission or at risk of transmission. The geostatistical map presents estimates of these serological indicators beyond the sampled locations, allowing evidence-based intervention targeting to take place beyond sampled clusters, along with estimates of model uncertainty demonstrating settings where further data may be needed for decision making.[60,61]

Should serology become a primary indicator for malaria surveillance, it may be worthwhile to review the recommended sampling strategy for serological indicators, to ensure a cost-efficient, timely and appropriately powered survey. Further developments to this work and exploration of the use of serological indicators as part of a package of surveillance tools in Ethiopia could be validation of measured seroprevalence and model predictions against other available data, including clinical burden recorded routinely at health facilities and cluster-level Malaria Indicator Survey data.

These data represent the spatial integration of simple survey design with a relatively basic laboratory assay that can subsequently guide malaria control and surveillance. The approach has particular utility in low-transmission settings and, therefore, has important applications for malaria elimination.

Authors' addresses: Ruth A. Ashton, Malaria Consortium, London, United Kingdom, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, E-mail: r.ashton@malariaconsortium.org. Takele Kefyalew, Malaria Consortium, Addis Ababa, Ethiopia, E-mail: t.kefyalew@malariaconsortium.org. Alison Rand, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, E-mail: alison.rand@pirbright.ac.uk. Heven Sime, Ethiopian Public Health Institute, Addis Ababa, Ethiopia, E-mail: hevensime@yahoo.com. Ashenafi Assefa, Ethiopian Public Health Institute, Addis Ababa, Ethiopia, E-mail: ashyaega@yahoo.com. Addis Mekasha, Oromia Regional Health Bureau, Addis Ababa, Ethiopia, E-mail: addismekasha@yahoo.com. Wasihun Edosa, Oromia Regional Health Bureau, Addis Ababa, Ethiopia, E-mail: wasihunchange@gmail.com. Gezahegn Tesfaye, Malaria Consortium, Addis Ababa, Ethiopia, E-mail: g_tesfaye@yahoo.com. Jorge Cano, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, E-mail: jcano.ortega@lshtm.ac.uk. Hiwot Teka, President's Malaria Initiative, United States Agency for International Development, Addis Ababa, Ethiopia, E-mail: hteka@usaid.gov. Richard Reithinger, President's Malaria Initiative, United States Agency for International Development, Addis Ababa, Ethiopia, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, and RTI International, Washington, DC, E-mail: rreithinger@yahoo.co.uk. Rachel L. Pullan, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, E-mail: rachel.pullan@lshtm.ac.uk. Chris J. Drakeley, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, E-mail: chris.drakeley@lshtm.ac.uk. Simon J. Brooker, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, E-mail: simon.brooker@lshtm.ac.uk.

REFERENCES

1. World Health Organization, 2012. *Disease Surveillance for Malaria Elimination*. Available at: http://whqlibdoc.who.int/publications/2012/9789241503334_eng.pdf. Accessed November 4, 2013.
2. Greenwood BM, 2008. Control to elimination: implications for malaria research. *Trends Parasitol 24:* 449–454.
3. Cotter C, Sturrock HJ, Hsiang MS, Liu J, Phillips AA, Hwang J, Gueye CS, Fullman N, Gosling RD, Feachem RG, 2013. The changing epidemiology of malaria elimination: new strategies for new challenges. *Lancet 382:* 900–911.
4. Tatem AJ, Smith DL, Gething PW, Kabaria CW, Snow RW, Hay SI, 2010. Ranking of elimination feasibility between malaria-endemic countries. *Lancet 376:* 1579–1591.
5. Okell LC, Bousema T, Griffin JT, Ouedraogo AL, Ghani AC, Drakeley CJ, 2012. Factors determining the occurrence of submicroscopic malaria infections and their relevance for control. *Nat Commun 3:* 1237.
6. Ishengoma DS, Francis F, Mmbando BP, Lusingu JP, Magistrado P, Alifrangis M, Theander TG, Bygbjerg IC, Lemnge MM, 2011. Accuracy of malaria rapid diagnostic tests in community studies and their impact on treatment of malaria in an area

7. Littrell M, Sow GD, Ngom A, Ba M, Mboup BM, Dieye Y, Mutombo B, Earle D, Steketee RW, 2013. Case investigation and reactive case detection for malaria elimination in northern Senegal. *Malar J 12:* 331.
8. Searle KM, Shields T, Hamapumbu H, Kobayashi T, Mharakurwa S, Thuma PE, Smith DL, Glass G, Moss WJ, 2013. Efficiency of household reactive case detection for malaria in rural southern Zambia: simulations based on cross-sectional surveys from two epidemiological settings. *PLoS ONE 8:* e70972.
9. Sturrock HJ, Novotny JM, Kunene S, Dlamini S, Zulu Z, Cohen JM, Hsiang MS, Greenhouse B, Gosling RD, 2013. Reactive case detection for malaria elimination: real-life experience from an ongoing program in Swaziland. *PLoS ONE 8:* e63830.
10. Roca-Feltrer A, Lalloo DG, Phiri K, Terlouw DJ, 2012. Rolling Malaria Indicator Surveys (rMIS): a potential district-level malaria monitoring and evaluation (M&E) tool for program managers. *Am J Trop Med Hyg 86:* 96–98.
11. Corran P, Coleman P, Riley E, Drakeley C, 2007. Serology: a robust indicator of malaria transmission intensity? *Trends Parasitol 23:* 575–582.
12. Corran PH, Cook J, Lynch C, Leendertse H, Manjurano A, Griffin J, Cox J, Abeku T, Bousema T, Ghani AC, Drakeley C, Riley E, 2008. Dried blood spots as a source of anti-malarial antibodies for epidemiological studies. *Malar J 7:* 195.
13. Stewart L, Gosling R, Griffin J, Gesase S, Campo J, Hashim R, Masika P, Mosha J, Bousema T, Shekalaghe S, Cook J, Corran P, Ghani A, Riley EM, Drakeley C, 2009. Rapid assessment of malaria transmission using age-specific sero-conversion rates. *PLoS ONE 4:* e6083.
14. Drakeley CJ, Corran PH, Coleman PG, Tongren JE, McDonald SLR, Carneiro I, Malima R, Lusingu J, Manjurano A, Nkya WMM, Lemnge MM, Cox J, Reyburn H, Riley EM, 2005. Estimating medium- and long-term trends in malaria transmission by using serological markers of malaria exposure. *Proc Natl Acad Sci USA 102:* 5108–5113.
15. Bousema T, Youssef RM, Cook J, Cox J, Alegana VA, Amran J, Noor AM, Snow RW, Drakeley C, 2010. Serologic markers for detecting malaria in areas of low endemicity, Somalia, 2008. *Emerg Infect Dis 16:* 392–399.
16. Oduro AR, Conway DJ, Schellenberg D, Satoguina J, Greenwood BM, Bojang KA, 2013. Seroepidemiological and parasitological evaluation of the heterogeneity of malaria infection in the Gambia. *Malar J 12:* 222.
17. Satoguina J, Walther B, Drakeley C, Nwakanma D, Oriero EC, Correa S, Corran P, Conway DJ, Walther M, 2009. Comparison of surveillance methods applied to a situation of low malaria prevalence at rural sites in The Gambia and Guinea Bissau. *Malar J 8:* 274.
18. Hsiang MS, Hwang J, Kunene S, Drakeley C, Kandula D, Novotny J, Parizo J, Jensen T, Tong M, Kemere J, Dlamini S, Moonen B, Angov E, Dutta S, Ockenhouse C, Dorsey G, Greenhouse B, 2012. Surveillance for malaria elimination in Swaziland: a national cross-sectional study using pooled PCR and serology. *PLoS ONE 7:* e29550.
19. Noor AM, Mohamed MB, Mugyenyi CK, Osman MA, Guessod HH, Kabaria CW, Ahmed IA, Nyonda M, Cook J, Drakeley CJ, Mackinnon MJ, Snow RW, 2011. Establishing the extent of malaria transmission and challenges facing pre-elimination in the Republic of Djibouti. *BMC Infect Dis 11:* 121.
20. Brooker S, Kolaczinski J, Gitonga C, Noor A, Snow R, 2009. The use of schools for malaria surveillance and programme evaluation in Africa. *Malar J 8:* 231.
21. Gitonga CW, Karanja PN, Kihara J, Mwanje M, Juma E, Snow RW, Noor AM, Brooker S, 2010. Implementing school malaria surveys in Kenya: towards a national surveillance system. *Malar J 9:* 306.
22. Gitonga CW, Kihara JH, Njenga SM, Awuondo K, Noor AM, Snow RW, Brooker SJ, 2012. Use of rapid diagnostic tests in malaria school surveys in Kenya: does their under-performance matter for planning malaria control? *Am J Trop Med Hyg 87:* 1004–1011.
23. Stevenson JC, Stresman GH, Gitonga CW, Gillig J, Owaga C, Marube E, Odongo W, Okoth A, China P, Oriango R, Brooker

SJ, Bousema T, Drakeley C, Cox J, 2013. Reliability of school surveys in estimating geographic variation in malaria transmission in the western Kenyan highlands. *PLoS ONE 8:* e77641.

24. Ghebreyesus TA, Haile M, Witten KH, Getachew A, Yohannes AM, Yohannes M, Teklehaimanot HD, Lindsay SW, Byass P, 1999. Incidence of malaria among children living near dams in northern Ethiopia: community based incidence survey. *BMJ 319:* 663–666.

25. Abeku T, Van Oortmarssen GJ, Borsboom G, De Vlas SJ, Habbema JDF, 2003. Spatial and temporal variations of malaria epidemic risk in Ethiopia: factors involved and implications. *Acta Trop 87:* 331–340.

26. Lautze J, McCartney M, Kirshen P, Olana D, Jayasinghe G, Spielman A, 2007. Effect of a large dam on malaria risk: the Koka reservoir in Ethiopia. *Trop Med Int Health 12:* 982–989.

27. Midekisa A, Senay G, Henebry GM, Semuniguse P, Wimberly MC, 2012. Remote sensing-based time series models for malaria early warning in the highlands of Ethiopia. *Malar J 11:* 165.

28. Graves PM, Richards FO, Ngondi J, Emerson PM, Shargie EB, Endeshaw T, Ceccato P, Ejigsemahu Y, Mosher AW, Hailemariam A, Zerihun M, Teferi T, Ayele B, Mesele A, Yohannes G, Tilahun A, Gebre T, 2009. Individual, household and environmental risk factors for malaria infection in Amhara, Oromia and SNNP regions of Ethiopia. *Trans R Soc Trop Med Hyg 103:* 1211–1220.

29. The Ethiopian Health and Nutrition Research Institute and partners, 2012. *Ethiopia National Malaria Indicator Survey 2011.* Addis Ababa, Ethiopia: Federal Ministry of Health.

30. Ashton RA, Kefyalew T, Tesfaye G, Pullan RL, Yadeta D, Reithinger R, Kolaczinski JH, Brooker S, 2011. School-based surveys of malaria in Oromia Regional State, Ethiopia: a rapid survey method for malaria in low transmission settings. *Malar J 10:* 25.

31. World Health Organization, 2001. *Iron Deficiency Anaemia: Assessment, Prevention and Control. A Guide for Programme Managers.* Geneva, Switzerland: World Health Organization.

32. Bousema T, Drakeley C, Gesase S, Hashim R, Magesa S, Mosha F, Otieno S, Carneiro I, Cox J, Msuya E, Kleinschmidt I, Maxwell C, Greenwood B, Riley E, Sauerwein R, Chandramohan D, Gosling R, 2010. Identification of hot spots of malaria transmission for targeted malaria control. *J Infect Dis 201:* 1764–1774.

33. Cook J, Kleinschmidt I, Schwabe C, Nseng G, Bousema T, Corran PH, Riley EM, Drakeley CJ, 2011. Serological markers suggest heterogeneity of effectiveness of malaria control interventions on Bioko Island, Equatorial Guinea. *PLoS ONE 6:* e25137.

34. Cook J, Reid H, Iavro J, Kuwahata M, Taleo G, Clements A, McCarthy J, Vallely A, Drakeley C, 2010. Using serological measures to monitor changes in malaria transmission in Vanuatu. *Malar J 9:* 169.

35. USGS-SRTM. *Shuttle Radar Topography Mission Digital Elevation Model.* Available at: http://srtm.csi.cgiar.org. Accessed November 4, 2013.

36. WorldClim. *WorldClim Global Climate Data.* Available at: http://www.worldclim.org. Accessed November 4, 2013.

37. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A, 2005. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol 25:* 1965–1978.

38. USGS. *SRTM Water Bodies Data Files.* Available at: https://lta.cr.usgs.gov/srtm_water_body_dataset. Accessed November 4, 2013.

39. Digital Chart of the World. Available at: http://diva-gis.org/gdata. Accessed November 4, 2013.

40. Matsuoka H, Nguon C, Kanbe T, Jalloh A, Sato H, Yoshida S, Hirai M, Arai M, Socheat D, Kawamoto F, 2005. Glucose-6-phosphate dehydrogenase (G6PD) mutations in Cambodia: G6PD Viangchan (871G > A) is the most common variant in the Cambodian population. *J Hum Genet 50:* 468–472.

41. SPOT Vegetation. *Spot 5 Vegetation.* Available at: http://www.spot-vegetation.com. Accessed November 4, 2013.

42. WorldPop Project. Available at: http://www.worldpop.org.uk/. Accessed November 4, 2013.

43. SEDAC. *Global Rural-Urban Mapping Project (GRUMP), v1.* Available at: http://sedac.ciesin.columbia.edu/data/collection/grump-v1. Accessed November 4, 2013.

44. Schwarz G, 1978. Estimating the dimensions of a model. *Ann Stat 6:* 461–464.

45. Akaike H, 1973. Information theory and an extension of the maximum likelihood principle. Petrov BN, Caski F, ed. *Proceeding of the Second International Symposium on Information Theory.* Budapest, Hungary: Akademiai Kiado, 276–281.

46. Diggle PJ, Ribeiro PJJ, 2007. *Model-Based Geostatistics.* New York, NY: Springer.

47. Spiegelhalter DJ, Best NG, Carlin BR, van der Linde A, 2002. Bayesian measures of model complexity and fit. *J R Stat Soc, B 64:* 583–616.

48. Swets JA, 1988. Measuring the accuracy of diagnostic systems. *Science 240:* 1285–1293.

49. Clements ACA, Lwambo NJS, Blair L, Nyandindi U, Kaatano G, Kinung'hi S, Webster JP, Fenwick A, Brooker S, 2006. Bayesian spatial analysis and disease mapping: tools to enhance planning and implementation of a schistosomiasis control programme in Tanzania. *Trop Med Int Health 11:* 490–503.

50. Fawcett T, 2006. An introduction to ROC analysis. *Pattern Recognit Lett 27:* 861–874.

51. Federal Democratic Republic of Ethiopia MoH, 2012. *National Malaria Guidelines*, 3rd Ed. Addis Ababa, Ethiopia: Ministry of Health (MoH).

52. Reid HL, Haque U, Roy S, Islam N, Clements AC, 2012. Characterizing the spatial and temporal variation of malaria incidence in Bangladesh, 2007. *Malar J 11:* 170.

53. Guerra CA, Howes RE, Patil AP, Gething PW, Van Boeckel TP, Temperley WH, Kabaria CW, Tatem AJ, Manh BH, Elyazar IRF, Baird JK, Snow RW, Hay SI, 2010. The international limits and population at risk of *Plasmodium vivax* transmission in 2009. *PLoS Negl Trop Dis 4:* e774.

54. Gething PW, Patil AP, Smith DL, Guerra CA, Elyazar IRF, Johnston GL, Tatem AJ, Hay SI, 2011. A new world malaria map: *Plasmodium falciparum* endemicity in 2010. *Malar J 10:* 378.

55. Malaria Atlas Project, 2013. *Malaria Atlas Project Maps of Malaria Risk in Ethiopia.* Available at: http://www.map.ox.ac.uk/explore/countries/ETH. Accessed September 23, 2013.

56. Diggle PJ, Menezes R, Su T, 2010. Geostatistical inference under preferential sampling. *Appl Stat 59:* 191–232.

57. Hartgers FC, Yazdanbakhsh M, 2006. Co-infection of helminths and malaria: modulation of the immune responses to malaria. *Parasite Immunol 28:* 497–506.

58. Garrett-Jones C, Shidrawi GR, 1969. Malaria vectorial capacity of a population of *Anopheles gambiae*: an exercise in epidemiological entomology. *Bull World Health Organ 40:* 531–545.

59. Poncon N, Tran A, Toty C, Luty AJ, Fontenille D, 2008. A quantitative risk assessment approach for mosquito-borne diseases: malaria re-emergence in southern France. *Malar J 7:* 147.

60. Oidtmann B, Peeler E, Lyngstad T, Brun E, Bang Jensen B, Stark KD, 2013. Risk-based methods for fish and terrestrial animal disease surveillance. *Prev Vet Med 112:* 13–26.

61. Rodriguez-Prieto V, Vicente-Rubiano M, Sanchez-Matamoros A, Rubio-Guerri C, Melero M, Martinez-Lopez B, Martinez-Aviles M, Hoinville L, Vergne T, Comin A, Schauer B, Dorea F, Pfeiffer DU, Sanchez-Vizcaino JM, 2014. Systematic review of surveillance systems and methods for early detection of exotic, new and re-emerging diseases in animal populations. *Epidemiol Infect 12:* 1–25.