SPATIO-TEMPORAL AND STATISTICAL ANALYSIS OF MALARIA IN THE AMANSIE WEST DISTRICT OF GHANA

Prosper, B.L.¹ and Duker, A.A.²

¹ Department of Environment and Resource Studies, University for Development Studies, Wa, Ghana ²Department of Geomatic Engineering, Kwame Nkrumah University of Science and Technology, Kumasi,

Ghana

ABSTRACT

In this study, Poisson variograms were used determine the spatial dependency of the risk of malaria which was then used to create surface maps from 2004 to 2009. Bayesian geostatitical approach was used to correlate the relationship between the elevation and the disease risk. Geographic Information System (GIS) was used to create the risk surfaces and overlays in the study. Buffered distances of 500, 1000, 1500 and 2000 m were used to overlay the disease risk map with forest, rivers/streams to find out its effects with the disease prevalence. The risk map created in this study showed an average of 20% rise yearly from 2004 to 2009. The results in the semi-variogram analysis with an average range of 2000 m showed that the disease incidence was local and not global. Areas which were more than 2km away from the water source (rivers/streams) recorded relatively higher cases except for some few within 1km of the Offin and Oda rivers. There was a varied effect of elevation with the disease prevalence was and a general trend of high incidence of the disease between 1 to 3 km from the edge of the forest.

Keywords: GIS, Geostatistics, Climatic Variables and Risk maps

INTRODUCTION

Malaria is a long life-threatening parasitic disease transmitted by female anopheles mosquitoes. This has contributed to child morbidity in the world. It threatens 2.4 billion people, or about 40% of the world's population living in the world's poorest countries (WHO, 2000). In semi-arid and highland regions of Africa, it is a major public health problem with over 200 million clinical episodes and nearly one million deaths occurring annually (WHO/UNICEF, 2005; Worall et al, 2004). This challenge cannot be allowed to go unnoticed since good health is not only a basic human need but also a fundamental human right and a prerequisite for economic growth (UN, 2003).

The malaria situation in Ghana is typical of sub-Saharan Africa making its transmission in southern Ghana an all- year -round affair and seasonal variation in the northern parts (Afari *et al.* 1995). The Ministry of Health (MOH) (2009) records shows that between 3-3.5 million cases of malaria are reported each year, over 900,000 of which are children under five years. However, the levels of malaria risk and transmission intensity exhibit significant spatial and temporal variability related to climate, altitude, topography, and human settlement pattern (Snow and Marsh, 2002; Abeku et al, 2003).

GIS is a computerized system utilized to process and manage spatial data. A GIS is capable of integrating topographical maps, satellite images, and aerial photos with attribute data such as demographic and socioeconomic characteristics and disease incidence. The systems have been used widely to produce maps of disease distribution and for analyzing spatial patterns in disease distribution (Cattani et al., 2001). In recent times, GIS and spatial analysis have proved equal to the task of improving mosquito monitoring as far as vector-borne diseases are concerned. It has therefore aided health professionals in their control and implementation strategies as far as malaria control is concerned (Griffith, 2005).

Topography or elevation variables have spatial variability which when analysed geostatistically could be used to determine the potential distribution of the vector and malaria risk areas. Poisson kriging offers more flexibility in modelling the spatial structure of disease risk and generates less smoothing, reducing the likelihood of missing areas of high risk as compared to other statistical methods as it inculcates the population size in computing the disease rates (Gooverts, 2005).

This paper therefore uses Poisson kriging and Bayesian approach to create risk maps and correlate the malaria risk with topographic and climatic covariates as well as the forest cover in the area to explore the spatial variability of the disease with these factors.

MATERIALS AND METHODS

STUDY AREA

The study was conducted in the Amansie West District, one of the twenty six (26) districts of the Ashanti Region of Ghana which has the highest situation of malaria. It is situated between latitudes 6° 00 N and 6° 45 N and longitudes 1° 30 W and 2° 15 W. The district has an area of 1,336 km² with the capital at Manso-Nkwanta. About 70% of the population are farmers whiles 22% are engaged in mining. The topography of the district is generally undulating with an average elevation of 210 m above mean sea level. It has a wide range of hills, which stretches across the north-western part of the district, especially around Manso-Nkwanta and Abore. These hills have an elevation of between 560m and 630m. The District is drained by the Offin and Oda rivers with an annual average rainfall of 1200mm. The District falls within the highest rainfall belt of Ghana having a double maximium rainfall pattern thereby supporting all year round farming activities. The temperature ranges from 22 °C to 30°C with a vegetation normally made up of secondary forests, thicket, and swamp and forb regrowth. The geology of the district is that of Proterozoic volcanic green stones with sedimentary rocks and granitoid intrusions (AWDA, 2004).

DIGITAL ELEVATION MODEL

Elevation data of the area were derived by creating a Digital Elevation Model (DEM) and creating a buffer of circles with radius 500m, 1,000m, 1,500m and 2,000m. The final elevations were then extracted using the mean of these buffer heights.

MALARIA PREVALENCE RATE

The malaria incidence per hundred (100) people of the population was then calculated i.e. Prevalence= {(number of cases/population)*100} for the communities under study. The malaria rates from 2004-2009 were thus computed as well as the total rates of the years under study.

BUFFER/OVERLAY OPERATIONS

With malaria areas closer to rivers/streams are seen to suffer most (Shilpa et al., 2004). A buffer distance of 500m, 1000m, 1500m and 2000m was generated since the 2000m, is not sufficient because of current ecological differences as far as mosquito flight distance is concerned. The 2000m is the average flight distance of the mosquito (Wim et al., 2003). The forest areas in the study area was classified using Erdas Imagine and buffered at a distance of 1km from the edge of the forest to the towns.

The combination of the risk i.e (elevation, rivers/streams and forest cover) were overlayed with the disease rates to investigate the combined influence of multiple factors on the disease prevalence. Assigned weights of 1, 3 and 5 generated the combined risk map using elevation, forest and rivers respectively according to their level of influence on the disease as suggested in other studies (Haque et al., 2009).

The Poisson risk values represented as points in terms of their relative risk sizes were then overlayed on the raster image obtained in buffering the forest, rivers/ streams and topography.

SPATIAL STATISTICAL ANALYSIS

Poisson kriging accounts for population rates in its modelling process as well as its generally flexible nature in the modelling and implementation of disease risk. The method is therefore utilized in this study.

We estimate the risk of contracting a disease at a given location m_{α} as $r(m_{\alpha})$ using a linear combination of K neighboring

data:
$$r_{PK}(m_{\alpha}) = \sum_{I=1}^{K} \lambda_i(m_{\alpha}) Z(m_i)$$
(1)

Where Z (m_i) is the rate observed at location m_i . The kriging weights are found by solving the following system of (k+1) linear equations:

$$\sum_{j=1}^{k} \lambda_{j}(m_{\alpha}) [C_{R}(m_{i}-m_{j})+\delta_{ij}\frac{d^{*}}{n(m_{i})}] + \mu(m_{\alpha}) = C_{R}(m_{i}-m_{\alpha}) \dots (2)$$

$$\sum_{j=1}^{k} \lambda_j(m_{\alpha}) = 1....(3)$$

Where $\frac{d^*}{n(m_i)}$ is an error variance term that represents the variability arising from population size and is derived directly under the Poisson model for the counts as used by (Gooverts, 2005)

where $\delta_{ij} = 1$ if $m_i = m_j$ and 0 otherwise, $n(m_i)$ and d^* is the population-weighted mean of the N rates calculated below:

The incorporation of this term for a zero distance $m_i = m_j$ leads one to assign smaller kriging weights to rates that are computed from smaller populations and deemed less reliable. The term

 $\mu(m_{\alpha}) \alpha$ is the Lagrange parameter which is the results from the minimization of the estimation

variance subject to the unbias constraint on the estimator.

In solving the kriging system in (2) above there is a need to have a model of the spatial covariance of the unknown risk, $C_R(h)$, or equivalently its semivariogram $\gamma R(\mathbf{h}) = C_R(o) - C_R(h)$. The experimental semivariogram of the risk as used by Monestiez et al is computed and shown below:

$$\gamma_{R}(\mathbf{h}) = \frac{1}{2\sum_{\alpha=1}^{N(h)} \frac{n(m_{\alpha})n(m_{\alpha}+h)}{n(m_{\alpha})+n(m_{\alpha}+h)}} \sum_{\alpha=1}^{N(h)} \left\{ \frac{n(m_{\alpha})n(m_{\alpha}+h)}{n(m_{\alpha})+n(m_{\alpha}+h)} [z(m_{\alpha})-z(m_{\alpha}+h)]^{2} - d^{*} \right\} \dots \dots \dots (5)$$

where N(h) is the number of communities separated by a vector (**h**). The different spatial increments $z(m_{\alpha}) - z(m_{\alpha} + h)]^2$ are weighted by a function of their respective population sizes, $\frac{n(m_{\alpha})n(m_{\alpha} + h)}{n(m_{\alpha}) + n(m_{\alpha} + h)}$), a term which is

inversely proportional to their standard deviation.

This therefore accounts for the accuracy and reliability of the data given the small standard deviation involved. As inferred from Gooverts (2005), A permissible model $\gamma_R(h)$, is then fitted to the experimental semivariogram in order to obtain the semivariogram, or covariance value, for any possible distance (h). The model follows the weighted least-square regression procedure.

BAYESIAN REGRESSION MODEL

The present study hypothesizes that the risk of malaria infection has a dynamic relationship with elevation. Here, the study adopts a non-linear nonparametric Bayesian modeling approach for the effect of elevation on the risk of malaria infection.

Consider the observations $(y_i, h_i), i = 1, ..., n$, with response y_i , and h_i the elevation in communities $s_i \in \{1, ..., S\}$. The study assumes that the response variable follows Gaussian distribution, i.e. $y_i | \eta_i, \sigma^2 \sim N(\eta_i, \sigma^2/c_i)$, with unknown mean η_i of a nonparametric *additive* model of the form:

$$\eta_i = f\left(h_i\right)$$

where f(h) is nonlinear smooth functions of h(s)

PRIOR ASSUMPTIONS

The unknown model parameters were estimated by a fully Bayesian approach. A Bayesian P (enalized)-splines (Stefan and Andreas, 2004) was used to model the unknown function f(h). This approach assumes that an unknown smooth function f of a covariate h can be approximated by a polynomial spline of degree l defined on a set of equally spaced knots $h^{\min} = \zeta_0 < \zeta_1 < \cdots < \zeta_{s-1} < \zeta_s = h^{\max}$ within the domain of h. Such a spline can be written in terms of a linear combination of d = s + l. B-spline basis functions B_m , i.e.

$$f(h) = \sum_{m=1}^{d} \xi_m \cdot B_m(x).$$

The B-splines form a local basis since the basic functions B_m are only positive within an area spanned by l+2 knots. This property is essential for the construction of the smoothness penalty for P-splines. The estimation of f(x) is thus reduced to the estimation of the vector of unknown regression coefficients $\xi = (\xi_1, ..., \xi_m)'$ from the data. An essential factor in the estimation procedure is the choice of the number of knots. We chose a moderately large number of equally spaced knots (20) to ensure enough flexibility to capture the variability of the data. In the Bayesian approach, penalized splines are introduced by replacing the difference penalties with their stochastic analogues, i.e., first or second order random walk priors for the regression coefficients. A first order random walk prior for equidistant knots is given by:

$$\xi_m = \xi_{m-1} + u_m, \qquad m = 2, \dots, d$$

and a second order random walk for equidistant knots by:

$$\xi_m = 2\xi_{m-1} - \xi_{m-2} + u_m, \qquad m = 3, ..., d$$

where the $u_m \sim N(0, \tau^2)$ are Gaussian errors. Diffuse priors $\xi_1 \propto const$, or ξ_1 and $\xi_2 \propto const$, are chosen as initial values, respectively. The joint distribution of the regression parameters ξ_m for a first order random walk is defined as:

$$\xi_m | \xi_{m-1} \sim N(\xi_{m-1}, \tau^2)$$

and a second order random walk is defined as:

$$\xi_m | \xi_{m-1}, \xi_{m-2} \sim N(2\xi_{m-1} - \xi_{m-2}, \tau^2).$$

The first order random walk induces a constant trend for the conditional expectation of ξ_m given ξ_{m-1} and a second order random walk results in linear trend depending on the two previous values ξ_{m-1} and $\xi_{j,m-2}$. The joint distribution of the regression parameters $\xi = (\xi_1, ..., \xi_m)'$ is computed as a product of the conditional densities defined by the random walk priors. The general form of the prior for ξ is a multivariate Gaussian distribution with density:

$$p(\xi|\tau^2) \propto \exp\left(-\frac{\xi' K \xi}{2\tau^2}\right)$$

where the precision matrix K acts as a penalty matrix that shrinks parameters towards zero, or penalizes too abrupt jumps between neighboring parameters. Since the penalty matrix K is rank deficient, i.e. $k = \operatorname{rank}(K) < \dim(\xi) = d$, it follows that the prior for $\xi | \tau^2$ is partially improper with Gaussian prior $\xi | \tau^2 \propto N(0; \tau^2 K^-)$, where K^- is a generalized inverse of K. The tradeoff between flexibility and smoothness is controlled by the variance parameter τ^2 . The larger the variance, the rougher is the estimated functions, and vice versa.

POSTERIOR ESTIMATION

Fully Bayesian inference is based on the posterior distribution of the unknown parameters. In this approach, samples are drawn from the full conditionals of the unknown parameters given the data through MCMC simulations. For simplicity, let β represent the unknown function to be evaluated (i.e., $\beta = (f(h))$ and τ the variance component; the posterior distribution then equals

$$p(\beta, \tau | y) \propto p(y|\beta) p(\beta|\tau) p(\tau)$$

where $p(y|\beta)$ is the likelihood function of the data given the parameters and p(.) represents the probability density function. Full conditional for the unknown function f(h(s)) is multivariate Gaussian and, as a consequence, a Gibbs sampler for MCMC simulation is employed. Cholesky decompositions for band matrices have been used to efficiently draw random samples from the full conditional. The model has been implemented in public domain software for Bayesian analysis, BayesX ver 2.0 (Stefan and Andreas, 2004).

RESULTS

The Malaria risk surface was created using the Poisson semivariogram model of the disease incidence. There was a steady rise in the general prevalence of the disease but a reduction in the key towns that were reporting in the previous years. A great improvement concerning risk was seen in 2007. There was an average risk occurrence in those areas except for areas around the district capital that was still reporting very high cases. The cummulative malaria risk was consistent with the general steady rise in the incidences of the years under study as evidenced in the ministry of health report 2010.



Fig 1: Variogram Analysis



Fig 2: Poisson Risk Maps of Malaria, 2006-2010.

EFFECT OF RIVERS/STREAMS ON MALARIA RISK

The results in Fig 2 showed areas within 2km away from the water source (rivers/streams) recording relatively higher cases except for some few areas within 1km just nearer the Offin and Oda rivers that recorded higher cases consistent with the Poisson risk maps.



Fig 3 : Distance Risk Map of Malaria in relation to rivers/streams

MALARIA AND FOREST COVER

The results in Fig 4 showed areas nearer the forest reporting higher cases though less significant .There is a general trend of high disease incidence between 1-3 km from the forest edge. Beyond 4km however there is varied disease incidence. Most settlements are not within the forest except for two towns Dominase and Abore.



Fig 4: Distance Risk Map of Malaria in relation to the Forest Cover

MALARIA AND TOPOGRAPHY

The results in Fig 4.4.1 showed the areas with very high elevations have no settlements. Settlements on the higher and lower elevations showed varied cases.



Fig 5: Digital Elevation Model (DEM) and malaria prevalence

TOTAL COVARIATES AND MALARIA

The combination of the risk: elevation, water and forest cover using assigned weights generated the combined risk as shown in Fig 5 below according to their level of influence on the disease. These results also showed a varied relationship. Fig 6 is the non-weighted overlays of the combined covariate.



Overlay of Total Covariates

Fig 6: Covariate Maps

DISCUSSIONS

The risk map created for the Amansie West District corresponded with the Ghana Health Service report that shows a great challenge in malaria control in the District. The disease prediction map was enhanced in the study because the data was smoothed using the Poisson statistical methods as it incorporated the population rates in the process. The risk map showed a high number of cases of malaria in most areas especially around the district capital. It can be seen that the prevalence of malaria tended to increase every year.

There was generally a fair increase in the malaria risk from the year 2004 to 2009. Malaria rose averagely by 16% from 2004 to 2007 and by 20% during the period 2008 to 2009. There were many areas reporting high risk and low risk respectively in the recurring years. Other high risk areas saw some gradual reduction especially in the other years (e.g 2006 and 2007). Millenium Village Project (MVP) started the malaria campaign programme in the district in 2006 and may have contributed a great deal to reduce prevalence only to see a rise in the ensuing years. In 2006 and 2007 the MVP gave free medications to the whole community. The free medications led to a lot of referrals to the hospital thereby increasing the number of reported cases. The campaign and distribution of insecticide bed nets especially at the MVP coverage area in 2006 and 2007 therefore saw a

reduction as evidenced in the Poisson map. The coverage of these bednets according to the MVP report increased from 13% to 59% by 2008 reducing the disease prevalence except for the district capital areas. Under reporting as a result of economic challenges and transport around those areas may have also increased the disease burden.

The distance to the rivers showed a varied relationship with the disease prevalence. Some areas nearer the water bodies rather showed a low prevalence as a result of the fact that the main rivers Oda and Offin is effectively being utilized and because the river flows fast, there is no stagnation of the water body. However in some communities water ponds are created close to the streams and rivers in order to preserve water during the dry season. These ponds created become stagnant points that enhance the breeding of mosquitoes that may increase the malaria prevalence at areas far off from the river/streams. Mining pits also nearer the river/streams may enhance mosquito breeding. The dug out pits from minnning left uncovered also become breeding grounds for the malaria vector to replicate. Furthermore some of the areas are marshy areas that leave a lot of water residue which serve suitably for the vector to thrive and therefore increases malaria rate.

The disease incidence was seen to be very high in the forest and forest fringes as compared to plains or non forested areas within 1km-3km. These areas as a result of their high humid conditions with corresponding higher rains as seen in Fig 3 enhances the ecology of the vector to thrive thereby increasing the disease prevalence. Mosquitoes in forested areas according to Afrane et al. (2006) live longer than those in the deforested area in both dry and rainy seasons in the highlands. They have shortened gonotrophic cycles thereby producing more eggs and reproducing faster. The forest habitats coupled with its humid conditions therefore support rapid multiplication of the vector. This may have resulted in the high cases in the forested areas. Most of the settlements are not within the forest except for the few (within the forest) that are not only influenced by the forest conditions but also on valley bottoms where the streams and rivers pass.

These environments often have areas where water could stagnate. Beyond 4km the disease prevalence could be attributable to other factors such as minning pits, untidy surroundings marshy areas and the non usage of mosquito treated nets and repellants. Higher elevation in general has long been recognized to be associated with malaria due to its association with cooler temperatures that slows the development of anopheline vectors and the Plasmodium parasites they transmit. Most of the hills in the study area have no settlements as they normally even have forests in those areas. Settlements around higher elevations however showed varied malaria cases as can be confirmed by the Bayesian regression models clearly showing that disease incidence is not homogeneous. In this case malaria risk displayed an alternating results with elevation.

The other areas on lower elevations may be closely related to the streams/swamps and forest impact. This may have been as a result of the terrain being suitable for water accumulation on valley bottoms. In those areas water is not washed away when it rains as in the mountainous or higher altitudes.

The fact that the elevation varied in some instances with the disease prevalence may have been as a result of the fact that the elevation differences was generally insignificant on the district spatial scale as compared to generally known above 2000m altitudes.

The combined covariate risk map shown in Fig 5 and 6 with weights revealed a weak correlation with the malaria risk. This is in conformity with the results of the semi –variogram analysis that showed that there was a weak spatial dependence on the malaria risk. It can therefore be deduced that different factors (rivers/streams, forest, elevation, temperature and rainfall) affect the disease risk differently.

This study suggests that population-based spatial and temporal analyses of initial surveillance data would be very helpful in managing malaria epidemiology, by highlighting when and where limited public health preventive medical resources should be concentrated.

The lower risk areas were seen to be the small settlements with better sanitary conditions probably due to their lower population. These areas may not even report to the hospital at all. It was also realized that most of the low risk areas made progress in the malaria campaign involving the distribution of Insecticide Treated Nets, drug supply and sensitization programmes by the MVP in the district. In the highland areas most farmers farm leafy crops such as cocoyam, plantain and pawpaw. On these crops water often collect for some time over their leaves. Such a condition could also enhance the growth of the vector.

The semi-variogram analysis confirmed that the malaria incidence is local and not global. The range of the semi-variograms shows the weak spatial relationship and dependencies on the malaria cases. It therefore shows that malaria is not as infectious as cholera and other communicable diseases. This results also highlights the fact that malaria occurrence within each town was not imported and that the local differences in topographic variables, rainfall, temperature, forest cover and mining pits may be the reason accounting for the small spatial dependencies in the malaria transmission.

CONCLUSIONS

Malaria maps have always being a precursor to monitoring, evaluation and interventions in epidemiological control. Identifying geographic risk factors, the spatial distribution and populations at risk are all critical steps towards the disease eradication. The risk map created with Poisson statistical methods showed areas at risk especially in the central portions of the district capital. It also showed an average of 20% rise yearly from 2004 to 2009. The results in the semi-variogram analysis with an average range of 2000m showed how the disease incidence was local and not global. The local nature of the disease occurrence gives credence to the fact that the covariates used which were rivers/streams, forest, temperature, rainfall and elevation had different and independent influence on the malaria prevalence. Areas nearer the rivers/streams more than 2km away from the water source (rivers/streams) recorded relatively higher cases except for some few areas within 1km just nearer the Offin and Oda rivers. Moreover there was a varied effect of elevation with the disease prevalence although there was an alternating relationship using the Bayesian regression model. This varied statistical relationship may have resulted from the small spatial scale of the district with elevation differences less than 50m. There was a general trend of high disease incidence between 1-3 km from the forest edge and different factors beyond 4km.

REFERENCES

- Abeku TA, de Vlas SJ, Borsboom GJJM, Tadege A, Gebreyesus Y, Gebreyohannes H, Alamirew D, Seifu A, Nagelkerke NJD, Habbema JDF., 2004. Effects of meteorological factors on epidemic malaria in Ethiopia: a statistical modeling approach based on theoretical reasoning. *Parasitology*, 128:585-593.
- 2. Afari EA, Appawu M, Dunyo S, Baffoe-Wilmot A & Nkrumah FK.,1995. Malaria infection, morbidity and transmission in two ecological zones in Southern Ghana. *African Journal of Health Sciences* 2, 312-315.
- 3. A.W.D.A (2004): *Demographic Characteristics and Infrastructural distribution* (Ghana: Amansie West District Profile 2004) pp. 12- 26

- 4. Cattani P, Jannin J, Lucas P. Sleeping sickness surveillance: an essential step towards elimination. *Trop Med Int Health.* 2001; 6:348–361.
- 5. Goovaerts P., 2005.Geostatistical analysis of disease data: estimation of cancer mortality risk from empirical frequencies using Poisson kriging. Int J Health Geogr, 4:31.
- 6. Griffith DA., 2005. A comparison of six analytical disease mapping techniques as applied to West Nile Virus in the coterminous United States. Int J Health Geographics. 4:18.
- 7. Haque U, Ahmed SM, Hossain S, Huda M, Hossain A, Alam MS, Mondal D, Khan WA, Khalequzzaman M, Haque

R.2009. Malaria prevalence in endemic districts of Bangladesh. PLoS ONE , 4:e6737.

- 8. Ministry of Health Report (MOH): Ghana 2001
- 9. Shilpa H, Penny M., Errol V., and Donald RR., 2004.Spatial correlations of mapped malaria rates with environmental factors in Belize, Central America. Int J Health Geogr. 2004; 3: 6.
- 10. Snow RW, Marsh K., 2002. The consequences of reducing transmission of Plasmodium falciparum in Africa. *Adv Parasitol*, 52:235-264.
- 11. Stefan L and Andreas B., 2004. Journal of Computational and Graphical Statistics Vol. 13, No. 1: 183-212.
- 12. UN.2003.United Nations Technical Health and Housing Report
- 13. WHO Expert Committee on Malaria Twentieth Report. World Health Organization: Geneva 2000.
- 14. WHO/UNICEF: *World malaria report 2005*, World Health Organization/UNICEF. Report series: WHO/HTM/MAL/2005.1102; 2005.
- 15. Wim van der Hoek, Flemming K, Priyanie HA, Devika P, Piyaratne MK, Felix P A., 2003. Towards a risk map of malaria for Sri Lanka: the importance of house location relative to vector breeding sites *.International Journal of Epidemiology*, Volume 32, Issue 2: 280-285.
- 16. Worrall ES, Basu XX, Hanson K., 2003. The relationship between socio-economic status and malaria: a review of literature.: Paper for the WHO/TDR Scientific Working Group on Malaria, Geneva, Switzerland, 24–27 March 2003.