

**MALARIA STRATIFICATION, CLIMATE AND EPIDEMIC EARLY WARNING
IN ERITREA**

PIETRO CECCATO, TEWOLDE GHEBREMESKEL, MALANDING JAITEH, PATRICIA M.
GRAVES, MARC LEVY, SHASHU GHEBRESELASSIE, ANDOM OGBAMARIAM,
ANTHONY G. BARNSTON, MICHAEL BELL, JOHN DEL CORRAL, STEPHEN J.
CONNOR, ISSAC FESSEHA, EUGENE P. BRANTLY AND MADELEINE C. THOMSON

International Research Institute for Climate & Society, The Earth Institute at Columbia University, Lamont Campus, Palisades, New York; National Malaria Control Program, Ministry of Health, Asmara, Eritrea; Center for International Earth Science Information Network (CIESIN), Columbia University, New York; EpiVec Consulting, Atlanta, Georgia; Department of Research and Human Resource Development, Ministry of Health, Asmara, Eritrea; Civil Aviation Authority, Meteorological Services, Asmara, Eritrea; RTI International, Washington DC.

ABSTRACT

Eritrea has a successful malaria control program but it is still susceptible to devastating malaria epidemics. Monthly data on clinical malaria cases from 242 health facilities in 58 *subzobas* (districts) of Eritrea from 1996 to 2003 were used in a novel stratification process using principal component analysis and non-hierarchical clustering to define five areas with distinct malaria intensity and seasonality patterns, to guide future interventions and development of an epidemic early warning system.

Relationships between monthly clinical malaria incidence by *subzoba* and monthly climate data from several sources, and with seasonal climate forecasts, were investigated. Remotely-sensed climate data were averaged over the same *subzoba* geographic administrative units as the malaria cases. Although correlation was good between malaria anomalies and actual rainfall from ground stations (lagged by two months), the stations did not have sufficiently even coverage to be widely useful. Satellite derived rainfall from the Climate Prediction Center Merged Analysis of Precipitation was correlated with malaria incidence anomalies, with a lead time of 2-3 months. NDVI anomalies were highly correlated with malaria incidence anomalies, particular in the semi-arid north of the country and along the northern Red Sea coast, which is a highly epidemic prone area.

Eritrea has two distinct rainy seasons in different parts of the country. The seasonal forecasting skill from Global Circulation Models for the June-July-August season was low except for the Eastern border. For the coastal October-November-December season, forecasting skill was good only during the 1997-1998 El Niño event. For epidemic control, shorter range warning based on remotely-sensed rainfall estimates and an enhanced epidemic early detection system based on data derived for this study are needed.

INTRODUCTION

Malaria Control in Eritrea. Eritrea is a malaria epidemic-prone country in the horn of Africa. It is divided into three main topographical regions (the western lowlands, central highlands, and eastern coastal lowlands) with different rainfall seasonality patterns. There are six administrative regions (zobas) with 58 districts (subzobas) (Figure 1) and approximately 1,500 villages. The estimated population is 3.5 million.

A survey in 2002 estimated malaria prevalence at 2%, with some villages having up to 30% prevalence.¹ Eritrea has had a successful malaria control program over the last 8 years since a devastating epidemic following the heavy rains associated with the 1997-1998 el Niño event.² An intensive mixture of interventions has been applied,² including widespread coverage with free impregnated mosquito nets to all malaria risk areas,³ indoor residual spraying with DDT and malathion in some areas, provision of prompt treatment by village health agents, source reduction and chemical or biological larval control.⁴

To improve epidemic control in climate sensitive regions, the World Health Organization (WHO) has proposed a framework for the development of integrated malaria early warning systems (MEWS),^{5, 6} based on vulnerability monitoring, seasonal climate forecasting, environmental and meteorological monitoring, and epidemiological surveillance. Here we explore the potential for using these indicators in the development of a malaria early warning system in Eritrea.

Vulnerability Assessment. Malaria risk maps derived from climate driven malaria transmission models have been used to indicate areas of stable and unstable malaria,⁷ and thereby indicate areas vulnerable to climate related epidemics. Early attempts to create a malaria risk map based on the modeled relationship between village-based prevalence data in Eritrea and a wide range of environmental and climatic predictors was found to be unsatisfactory.^{8, 9} In this paper we report on a new methodology based on routine incidence data, designed to indicate the intensity and seasonality of transmission at the district level.

Climate and environmental monitoring. In most of Eritrea, the majority of the rainfall occurs during July-August. In the eastern portion of the country, where a steep escarpment degrades towards the Red Sea, rainfall is considerably lower, and peaks in October-January. Both ground station measures and satellite-derived estimates of rainfall are potential proxies for use by malaria control staff. Comparisons of the two sources in neighboring Ethiopia indicated that certain satellite-derived rainfall estimates (RFE) products may provide sufficiently accurate rainfall information in the region to be used routinely by malaria control.¹⁰

There is a strong relationship between normalized difference vegetation index (NDVI), a measure of environmental ‘greenness’, and rainfall both spatially and temporally over East Africa, as well a demonstrated relationship between NDVI, rainfall, and the El Niño-Southern oscillation (ENSO).^{11 - 13}

Seasonal climate forecasts. For certain geographic areas and seasons, climate forecasts offer a degree of predictability of climate fluctuations at a seasonal (i.e., ~ 3 month) lead time.¹⁴ Predictability is highly dependent on the extent to which regional climate for a given season is determined by sea surface temperature (SST) patterns of global oceans; in particular SSTs in the tropical Pacific (i.e., the El Niño-Southern Oscillation, ENSO). There is some indication that the summer rainfall in Eritrea is related to ENSO¹⁵ and therefore potentially predictable.

Significant correlations between malaria incidence and observed SSTs (including time-lagged SSTs) have been observed.^{16, 17} Building on these observations, seasonal climate forecasts produced from the real-time operational multi-model ensemble DEMETER forecast systems have started to contribute to the development of malaria early warning in Botswana.^{18, 19}

Surveillance Data. In many African countries malaria data are collected at health facilities but often not looked at until they are summarized in the form of routine annual reports. Recently, many countries including Eritrea have attempted to improve the collection and use of health facility data, by updating health management information systems (HMIS). The aim is to improve estimates of the burden of disease at the national level while enhancing quality of care,

accountability and decision making at the local level. Weekly surveillance data is particularly valuable for the detection of unusual situations, possibly an epidemic, which may require a prompt response.²⁰

Unfortunately, to date, there is very little information which feeds back to the district level for improved decision making, thus undermining the incentives for improved reporting.²¹ Furthermore, biases in data reported which result from quality of service, drug availability, physical and social access, and health seeking behavior feed the widespread distrust of passive surveillance data collection as a tool in routine malaria decision making.²² Furthermore, for clinically diagnosed malaria cases (as opposed to those confirmed by laboratory diagnosis) there is often considerable overestimation of the case numbers as well as delay in appropriate treatment for non-malaria fever patients²³. Despite these limitations, many countries have no alternative but to make best use of the information available through passive health surveillance systems. Geographical and seasonal patterns of malaria transmission are usually clearly evident from such data.²⁴

In Botswana, where the HMIS is not only functioning well but where all suspected malaria cases are routinely laboratory-confirmed, a strong correlation is observed between clinical and confirmed cases¹⁷. Thus in epidemic-prone areas, routine surveillance of clinical cases may be useful for determining past thresholds and for the early detection of epidemics. A recent study from Kenya has shown that the quality of routinely collected health surveillance data from the HMIS can be dramatically improved through statistical methodologies which impute missing data using geospatial techniques²⁵.

METHODS

Surveillance data. Eritrea's National Health Management Information System (NHMIS) was revised in 1998 with support from USAID and John Snow Inc. This revision involved computerization of a previously existing paper system for monthly reporting from each health facility, standardized coding of case definitions, coding of health facilities and administrative areas, reduction in numbers of age groups reported, and establishment of reporting timelines. There were no major changes in malaria case definitions or unit of reporting (health facility).

For this study, the numbers of outpatient clinical malaria cases were extracted from the NHMIS Access database by month and health facility for the years 1998 to 2003. Data for 1996 to 1997 were obtained from similar Malaria Control Programme records stored in Excel by health facility and month.

The NHMIS listed 325 health facilities in 1998. We restricted this to 242 sites by exclusion of private doctors, worksite clinics, national referral hospitals in Asmara, and exclusion of 3 non-functioning health facilities. The number of health facilities in the system had increased to 383 by 2003, of which many were new small private clinics as well as one new

National Referral Hospital. In this analysis, we use only data for the original 242 selected facilities. Administrative (*subzoba* and *zoba*) boundary files were obtained from the National Statistics and Evaluation office, Asmara, Eritrea, and visualized in ArcView 3.3. Malaria cases were summed over both age-groups (under and over 5 years) by *subzoba* and month (1-9 facilities per *subzoba*, average 4.3). Malaria incidence per 1000 persons per month by *subzoba* was estimated using the year 2000 population estimates in the NHMIS.

Imputation of missing values. If complete, the 8 years (1996-2003) of monthly incidence data for 58 *subzobas* would have records for 5568 *subzoba*-months. There were 191 missing data points (*subzoba*-months), representing 3.4% of the total matrix, impeding the use of principal component analysis which is dependent on complete data sets. Missing values were imputed through a sequence of multivariate stepwise regressions which estimated the variables with the smallest number of missing values first and proceeding until all missing values were estimated.²⁶⁻²⁸

Stratification map using NHMIS data. We used the ADDATI software developed for exploratory data analysis, including principal components and clustering analysis²⁹. Monthly malaria incidence estimates, computed as arithmetic average of values in that month during the 8-year period for each of the 58 *subzobas*, were used for the analysis. The processing sequence in ADDATI includes i) principal component analysis followed by ii) non-hierarchical clustering adapted from Diday³⁰ on the most significant principal components.

The principal component analysis (PCA) is a technique for simplifying a data set, by reducing multidimensional data sets to lower dimensions in the data³¹. The reduction in dimensions of the data set is necessary to run and accelerate the clustering analysis which is a classification of objects into different groups such that each group shares a common trait. A total of 12 principal components were generated from the monthly average incidence data. Using eigenvalue and proportion of variance explained by each component criteria, we retained, for further analysis, the 6 first principal components which explained 99.37% of the cumulative variance in the data.

The iterative non-hierarchical clustering method groups the *subzobas* into a specific user-defined number of clusters. ADDATI generates the cluster centers and allocates the *subzobas* to the clusters based on proximity to initial centers (“seeds”).²⁹ The data were initially clustered into 2, 3, 4, 5, 6, 8 and 12 clusters for analysis. Clusters were visualized in ESRI ArcView 3.3™.

Climate datasets. While there is a variety of rainfall datasets for Eritrea, few are suitable for analyses of long-term trends and variability. After investigating a range of datasets, we concluded that three were useful for assessing climate variability in Eritrea during the study period (Table 1): Rain gauge data from meteorological stations (locations shown in Figure 2); merged gauge and satellite rainfall estimates (CMAP); and vegetation index (NDVI) from NOAA AVHRR computed as a monthly maximum value composite. A fourth potential dataset (from the Climate Research Unit at the University of East Anglia, UK) consisted of interpolated meteorological station data, but was only available up until 1998. Using the first three sources, the monthly anomalies (departure from long term mean) of climate environmental variables were

calculated. The monthly anomalies, both for climate environmental variables and for incidences data, are used for the analysis. Anomalies are used instead of recorded values because both climate variables and malaria incidences have a similar seasonal pattern. Statistical analysis of recorded values would indicate a very high correlation between climate and malaria incidences only due to the fact that the temporal evolution of climate variables and malaria incidences follow the same seasonality. In order to eliminate the seasonal cycle effect and study only the impact of variations in climate with variations in malaria incidence, the long-term mean is subtracted from the recorded data.

The satellite-derived climate variables from the IRI data library were averaged over the same administrative boundaries as the incidence data, using *subzoba* boundaries from the National Statistics and Evaluation Office, Eritrea. Whilst it would have been preferable to use smaller boundaries such as health facility catchment areas, these are not yet available.

Seasonal climate forecasts. We used rainfall output from a global circulation model (GCM) ECHAM 45 (Table 1) to predict September NDVI using forecasted June-July-August (JJA) rainfall, and to predict January NDVI using forecasted October-November-December (OND) rainfall.

RESULTS

Reporting rate. The average reporting rate by month at the health facility level over the whole 8 year period was 76.9%. Of the 242 health facilities, the majority (155 (64.0%)) provided over 80% of required reports, 38 facilities (15.7%) provided 60 to 80% of reports, 20 (8.3%) provided 40 to 60% of reports, 19 (7.9%) provided 20 to 40% of reports and only 10 (4.1%) provided under 20% of the required reports. In some cases this is because the facility was not functioning for part of the time period, but detailed information on this is not available.

The estimates of reporting completeness by year were as follows:

1996: 52.1% of reports received, 59.9% of facilities reporting;
1997: 71.2% of reports received, 74.4% of facilities reporting;
1998: 78.1% of reports received, 84.7% of facilities reporting;
1999: 83.3% of reports received, 87.2% of facilities reporting;
2000: 77.4% of reports received, 92.1% of facilities reporting;
2001: 84.8% of reports received, 93.8% of facilities reporting;
2002: 86.4% of reports received, 92.6% of facilities reporting;
2003: 89.5% of reports received, 93.0% of facilities reporting.

As mentioned in the Methods section, there were a potential 5568 data points (*subzoba*-months) from the 58 *subzobas* and 96 months of study. There were 191 missing data points where no reports were received in a particular *subzoba*, representing 3.4% of the total matrix.

Vulnerability mapping with surveillance data. The 5-cluster map (Figure 3) was found to be most representative of what is expected in the field and is consistent with expert opinion.

With five clusters we were able to capture differences in both intensity and the seasonal dynamics of malaria incidence. This is particularly important in Eritrea as the coastal region is dominated by October-November-December (OND) rains as opposed to July-August-September (JAS) rains for the rest of the country. However this analysis was able to differentiate distinct risk regions (groups of *subzobas*) within the Red Sea coastal region of Eritrea as well as different intensities of transmission within the areas susceptible to the OND rainfall season. One can argue that the difference in intensity (or malaria incidences) observed between the clusters could be related to the number of health facilities per subzoba which could affect the coverage and hence the number of malaria cases reported from a subzobas. We analyzed the number of health facilities and number of reports received per year to verify whether this hypothesis could influence the intensity. Tables 2 and 3 show that there is indeed a difference in number of health facilities between clusters, but the average number of reports received per cluster and per year is very similar and does not influence the number of incident cases recorded (*e.g.* Cluster 2 has the highest number of health facilities and still is classified as low incidence rate).

In addition to clarifying the seasonal pattern and intensity of transmission in different areas, the stratification map produced is useful for targeting and timing interventions such as indoor residual spraying to the highest risk areas.

Climate/environmental predictors of malaria incidence. The relationship of monthly malaria incidence anomalies to anomalies in climate/ environmental data (and each anomaly value squared) for each *subzoba* was assessed using Spearman's and Pearson's rank correlation.

There is a good correlation between rainfall and malaria anomalies in some of the *subzobas* where rain gauge station data are available (22 out of 58 *subzobas*) with a lead-time of 2 months (Figure 4). However, there is no information for large tracts of the country, suggesting that while this data source may be used in specific districts it is not suitable for national coverage.

Satellite derived estimates of rainfall. Rainfall estimate anomalies were also moderately correlated with malaria incidence anomalies with an average lead-time of 2-3 months. Although the strength of the relationship was less than that of the gauge data, this data source had the advantage of complete coverage of the country.

Satellite derived estimates of vegetation 'greenness'. NDVI anomalies were correlated with malaria incidence anomalies, particularly in the semi-arid north of the country and along the coast in Northern Red Sea (Figure 5). Areas of high correlation are largely those identified as 'sparse vegetation' from the Africover land cover map produced by the Food and Agriculture Organization of the United Nations³². An advantage of NDVI over rainfall is that it is indicative of soil moisture and is therefore capable of picking out areas where rainfall runoff may have a significant impact on malaria transmission.

Seasonal climate forecasts. We retrospectively forecast June-August rainfall data using a general circulation model forced with May SST anomalies and correlated the results with September NDVI over the Greater Horn of Africa including Eritrea. This forecast would normally be available towards the end of June. The skill is good over the northern part and eastern coast of the Greater Horn of Africa with correlation ranging from 0.4 to 0.7 (not shown). However, over Eritrea the area of predictability was limited to the eastern border and there is low skill far from the sea (north western region).

Results show that good correlations (with coefficients between 0.6-0.8) are obtained in Eritrea for the period October-November-December (OND). However, correlation values dropped dramatically if the year 1997 is not taken into account, implying that good predictability in terms of rainfall could be achieved only during a very strong el Niño event.

DISCUSSION

Malaria stratification is a classification of areas according to the risk of malaria. It is a way to set priorities and target prevention efforts to the areas where they are most needed. It can highlight areas where the control program needs extra effort, and helps to make the best use of resources. Changes in stratification over time can also help to indicate the areas at most risk of epidemics.

Since control activities are mostly organized by administrative unit such as *subzoba*, a classification of these units is more useful to the program than a strict ecological stratification, where different environmental strata may overlap more than one *subzoba*. In future, risk stratification can be extended to smaller units within *subzobas* if desired. The availability from the International Research Institute for Climate and Society (IRI) of remotely sensed data averaged over the *subzoba* administrative unit, together with malaria cases in the same geographic units, was highly beneficial to the stratification process.

Our results show the ADDATI software, freely available for agro-climatic applications, can be a useful tool for malaria stratification. Although the program was intended for classification of satellite-derived time series land cover data into areas of similar spatio-temporal patterns in vegetation development, it is relatively easy to use for clustering districts with similar malaria incidence and seasonality.

Using only health-facility clinical malaria incidence data aggregated over administrative divisions we identified 5 clusters (groups of administrative units) of malaria risk based on both intensity and dynamics. The cluster classification provides guidance for targeting extra malaria control methods (such as indoor residual spraying) to the highest incidence areas. It also clarifies the seasonality differences between regions and indicates optimum timing of interventions.

Stratification of malarious areas by average parasite incidence has been applied previously^{33, 34}. Our current methodology is an advance because it identifies not only geographic areas with similar malaria intensity but also areas with similar temporal dynamics. It also relies

only on readily available surveillance data on malaria incidence to stratify in areas with significant seasonal variation in incidence. Given that many malarious countries have established or improved their health management information systems as repositories for surveillance and monitoring data, development of a stratification method that depends on such data has great potential.

While the problem of underreporting of disease incidence from underserved areas is always present, this may be compensated by the availability of much larger and more detailed datasets than are usually available from cross-sectional prevalence surveys. Problems of inconsistent reporting can be minimized by restricting analysis only to health facilities with reliable reporting, ideally of confirmed malaria cases.

The dataset used for stratification also provides necessary outcome data for assessing the effectiveness of different control methods used in the country over the last 8 years. However, since malaria anomalies and rainfall/NDVI are clearly associated in Eritrea it is imperative that an analysis of the impact of interventions take into account climate/environment variability.

The routine data is also a rich source of past information at the *subzoba* and health facility level for generating epidemic threshold values for early detection of abnormal numbers of cases. This task is currently in progress.

The climate/environmental variables used in this report can be routinely monitored, and in fact some are already used within Eritrea for Food Security and Locust monitoring. These environmental variables have the potential to be partially predicted using seasonal climate forecasts, although this predictability is unlikely to be as good as other parts of east and southern Africa.^{35, 36}

There is a known trade-off between timing and accuracy in any such warning system. In our example:

- a. Health surveillance using epidemic thresholds detects the early phase of epidemics and provides a couple of weeks warning of epidemic peak
- b. Environmental monitoring of NDVI provides concurrent prediction of malaria anomalies
- c. Rainfall from gauge data provides prediction of malaria anomalies with a lead time of 2-3 months for some *subzobas* with meteorological stations
- d. Rainfall from merged gauge and satellite data (CMAP) predicts September or January NDVI (when peak malaria occurs) with 2-3 months lead time
- e. General circulation models predict rainfall with a lead-time of 2 months and NDVI with a lead-time of 4 months, but with reasonable skill only in el Niño years for most parts of Eritrea.

Stratification and definition of the relationship between climate and malaria are significant steps towards development of an early warning system. Once a picture of the 'expected' number of monthly (or perhaps in future weekly) malaria cases in a *subzoba* or cluster has been built up, together with the expected climate variables in the same areas, all the necessary components of an early warning system are in place.

AUTHORS' ADDRESSES

Anthony Barnston, Pietro Ceccato, Stephen J. Connor, John del Corral, Madeleine C. Thomson: International Research Institute for Climate and Society (IRI), The Earth Institute at Columbia University, Lamont Campus, 61 Route 9W, Palisades, New York 10964-8000. Tel (++)1 845 680 4413, Fax (++) 1 845 680 4864, Email: pceccato@iri.columbia.edu .

Tewolde Gebremeskel: National Malaria Control Program, Ministry of Health, Asmara, Eritrea

Malanding Jaiteh, Marc Levy: Center for International Earth Science Information Network (CIESIN), Columbia University, New York

Patricia M. Graves: EpiVec Consulting, 606 Kimberly Lane NE, Atlanta GA 30306.

Tel: 404 293 3529; Fax 770 488 4258. email : epivec@comcast.net

Andom Ogbarmariam, Shashu Ghebreselassie: Dept of Research and Human Resources, Ministry of Health, Asmara, Eritrea

Issac Fesseha: Civil Aviation Authority, Meteorological Services, Asmara, Eritrea

Eugene Brantly, RTI International, One Metro Center, 701 13th Street, NW Suite 750, Washington, DC 20005.

ACKNOWLEDGEMENTS

We thank the Minister of Health for Eritrea, Mr Saleh Meko, the Director General for Health, Mr Berhane Tensae, and the head of Disease Prevention and Control, Dr Goitom Mebrahtu, for their support. We are very grateful to Kiros Sereke, Afwerki Araia, Andemariam WeldeMicael, Solomon Neguse, Meles GhebreYesus, Asrat GhebreLul, Mehari Zerom, Asmelash Gegziabher, Helen Fekadu, Solomon Mengistu, Yohannes Bein, David Sintasath, and Dr. Josephat Shililu of the NMCP and Ezra Kidane, Amanuel Kifle and Samuel Goitom from the NHMIS. We thank Matthew Lynch of USAID, Lisa Lukang of CIESIN and Tsegay Wolde-Georgis of IRI for input and support . This study was funded by the USAID Environmental Health Project, Contract HRN-1-00-99-00011-00 sponsored by Office of Health, Infectious Disease and Nutrition, Bureau for Global Health, USAID, Washington DC 20523. We thank the National Statistics and Evaluation Office, Asmara, Eritrea and the Water Resources Department, Ministry of Lands, Water and Environment for provision of GIS data. The following organizations supported the malaria control program of Eritrea during the period of study: The World Bank HAMSET project; USAID; WHO; Italian Cooperation; UNICEF.

REFERENCES

1. Sintasath DM, Ghebremeskel T, Lynch M, Kleinau E, Bretas G et al (2005). Malaria prevalence and associated risk factors in Eritrea. *Am J Trop Med Hyg* 72: 682-687.
2. Nyarango PM, Gebremeskel T, Mebrahtu G et al (2006). A steep decline in malaria morbidity and mortality trends in Eritrea between 2000 and 2004: the effect of a combination of control methods. *Malaria J* 5:33.
3. Macintyre, K, Keating J, Okbaldt YB, Zerom M, Sosler S, Ghebremeskel T, Eisele TP (2006). Rolling out insecticide treated nets in Eritrea: examining the determinants of possession and use in malarious zones during the rainy season. *Trop Med Int Health* 11(6): 824-33.
4. Shililu JI, Tewolde GM, Brantly E, Githure JI, Mbogo CM, Beier et al (2003) Efficacy of *Bacillus thuringiensis israelensis*, *Bacillus sphaericus* and temephos for managing *Anopheles* larvae in Eritrea. *J Am Mosq Cont Assoc* 19: 251-258.
5. WHO (2001). Malaria Early Warning Systems, Concepts, Indicators and Partners. A framework for field research in Africa. Geneva: WHO.
6. WHO (2004). *Field Guide for Malaria Epidemic Assessment and Reporting. Draft for field testing*. Geneva: World Health Organization; 2004. WHO/HTM/MAL/2004.1097.
7. Craig MH, Snow RW, le Sueur D (1999). A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitol Today* 15: 105-111.
8. Bretas G (2001). Malaria Risk Stratification in Eritrea. Report to the Environmental Health Project, USAID
9. Thomson M, Graves PM, Barnston AG, Bell M, Ceccato P, Connor S, del Corral J, Giannini A, Obsomer V, Wolde-Georgis T, Jaiteh M, Levy M, Lukang L. (2005). Towards a malaria early warning system for Eritrea. Final report to Environmental Health Project, USAID.
10. Dinku T, Ceccato P, Grover-Kopec E, Lemma M, Connor SJ, and Ropelewski CF, (in press). Validation and inter-comparison of satellite rainfall products over East African complex topography. *International Journal of Remote Sensing*
11. Davenport ML, Nicholson SE (1993). On the relation between rainfall and the normalized difference vegetation index for diverse vegetation types in east Africa. *International Journal of Remote Sensing* 14: 2369-2389.
12. Nicholson SE, Farrar TJ (1994). The influence of soil type on the relationships between NDVI, rainfall, and soil moisture in semiarid Botswana .1. NDVI response to rainfall. *Remote Sensing of Environment* 50: 107-120.
13. Anyamba A, Eastman JR (1996). Interannual variability of NDVI over Africa and its relation to El Nino Southern Oscillation. *International Journal of Remote Sensing* 17: 2533-2548.
14. Goddard L, Mason SJ, Zebiak SE, Ropelewski CF, Basher R, Cane MA, (2001). Current approaches to seasonal-to-inter-annual climate predictions. *International Journal of Climatology* 21: 1111-1152.
15. Seleshi Y, Demaree GR, (1995). Rainfall variability in the Ethiopian and Eritrean highlands and its links with the Southern Oscillation Index. *Journal of Biogeography*, 22(4-5): 945-952.
16. Kovats RS, Bouma MJ, Hajat S, Worrall E, Haines A (2003). El Nino and health. *Lancet* 362: 1481-1489.

17. Thomson MC, Connor SJ, Phindela T, Mason SJ (2005). Rainfall and sea-surface temperature monitoring for malaria early warning in Botswana. *Am J Trop Med Hyg* 73: 214-221.
18. Thomson MC, Doblaz-Reyes FJ, Mason SJ et al (2006). Malaria early warnings based on seasonal climate forecasts from multi-model ensembles. *Nature* 439: 576-579.
19. DaSilva J, Garanganga B, Teveredzi V, Marx S, Mason SJ, Connor SJ (2004). Improving epidemic malaria planning, preparedness and response in southern Africa. *Malaria J* 3: 37.
20. Allard R, (1998). Use of time-series analysis in infectious disease surveillance. *Bull WHO* 76(4): 327-333.
21. Chaulagai CN, Moyo Cm, Koot J, Moyo HB, Sambakunsi TC, Khunga FM, Naphini PD (2005). Design and implementation of a health management information system in Malawi: issues, innovations and results. *Health Policy Plan* 20(6): 375-84.
22. Chilundo B, Sundby J, Aanestad M. (2004). Analysing the quality of routine malaria data in Mozambique. *Malaria J* 3: 3.
23. Chandramohan D, Jaffar, S, Greenwood B. 2002. Use of clinical algorithms for diagnosing malaria. *Trop Med and Int Health* 7(1): 45-52.
24. Mabaso MLH, Craig M, Vounatsou P, Smith T. 2005. Towards empirical description of malaria seasonality in southern Africa: the example of Zimbabwe. *Trop Med and Int Health* 10(9): 909-918
25. Gething PM, Noor AM, Gikandi PW, Ogara EAA, Hay SI, Nixon MS, SnowRW, Atkinson PM. (2006). Improving imperfect data from health management information systems in Africa using space–time geostatistics. *PLOS Med* 3: 6.
26. Raghunathan TE, Lepkowski JE, Solenberger PW, Van Hoewyk JH (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodol* 27: 85.
27. Rubin DB (1996). Multiple imputation after 18+ years. *J Am Stat Assoc* 91: (434): 473-489.
28. Schafer JL (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.
29. Griguolo S, Mazzanti M (1998). *ADDATI: un package per l'analisi esplorativa dei dati*. V.4.0. Guida all'uso, Padova, Editrice Libreria Progretto, 1998, 202 pp .
30. Diday, E. (1971) *Rev. Stat. Appl.* **19**, 19-34
31. Jolliffe, A.T. (1986). *Principal Component Analysis*. Springer, New York
32. Di Gregorio, A, Mansen LJM, (2000). Land cover classification system, classification concepts and user manual. Food and Agriculture Organization of the United Nations, Rome, 179 pp
33. Booman M, Durrheim DN, La Grange K, Martin C, Mabuza AM, Zitha A, Mbokazi FM, Fraser C, Sharp BL. (2000). Using a geographical information system to plan a malaria control program in South Africa. *Bull WHO* 78: (12): 1438-1444.
34. Wickramasinghe AR, Gunawardena DM, Mahawithanage ST. (2002). Use of routinely collected past surveillance data in identifying and mapping high risk areas in a malaria endemic area of Sri Lanka. *SE Asian Journal of Tropical Medicine and Public Health*. 33(4): 678-84.
35. Goddard L and Graham NE. (1999). The importance of the Indian Ocean for simulating precipitation anomalies over eastern and southern Africa. *J. Geophys. Res.* 104: 19099-19116.

36. Korecha D and Barnston AG. In press. Predictability of June to September rainfall in Ethiopia. *Mon. Wea. Rev.* 132:
37. New M, Hulme M, Jones P, (2000). Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate. *Journal of Climate* 13: 2217-2238.

Fig 2: Rainfall station locations in Eritrea (23 stations)

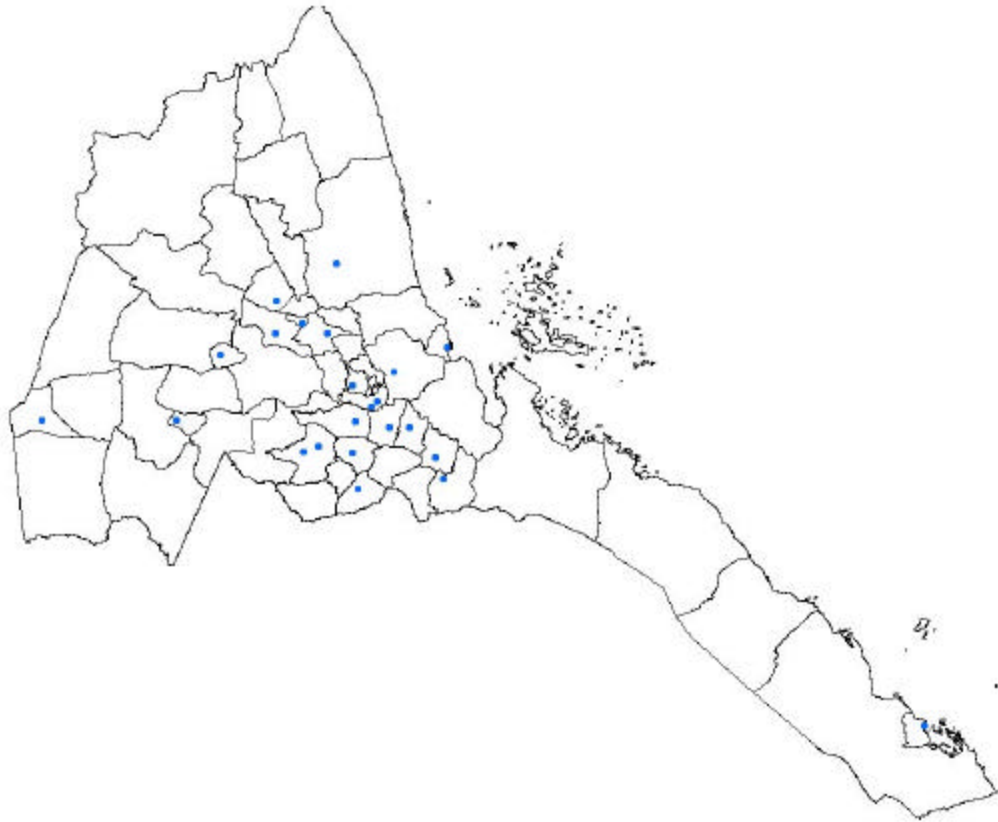


Figure 3: Stratification map of Eritrea malaria risk based on monthly clinical malaria data derived using principal components analysis and non-hierarchical clustering.

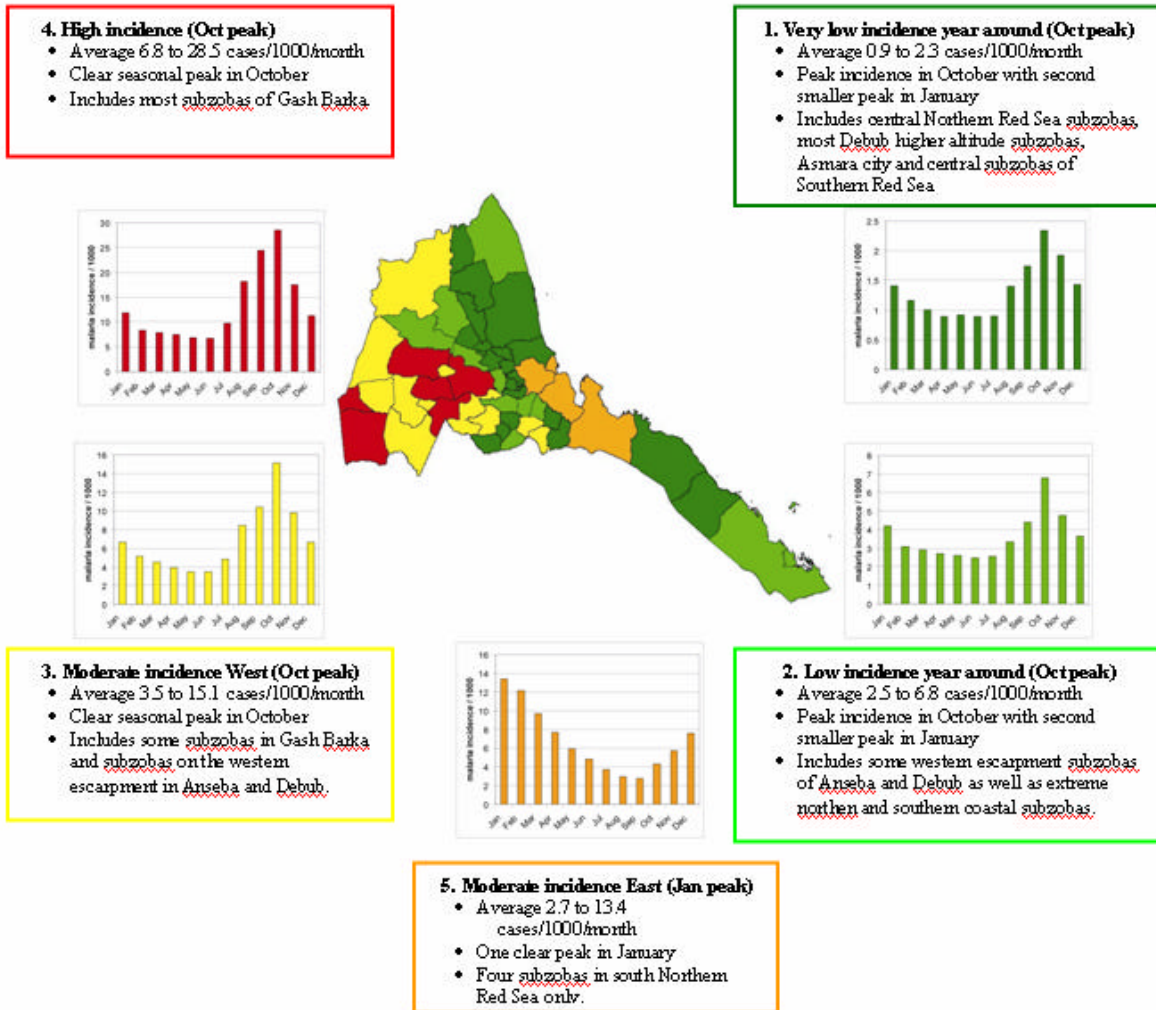


Fig 4 Variance in malaria incidence predicted by rainfall from meteorological stations lagged 2 months. The coefficient of determination (R^2) values indicates the proportion of variability in the malaria anomalies that is accounted for by the rainfall.

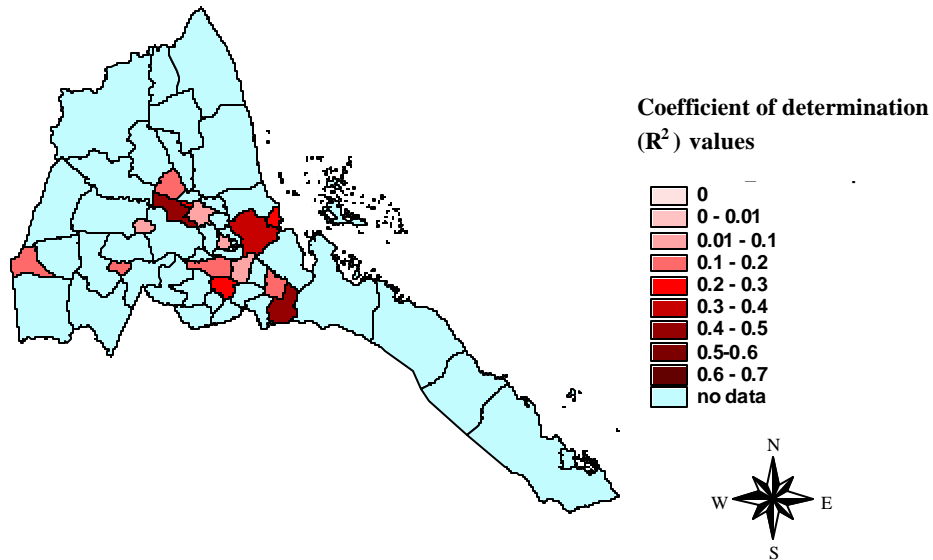


Fig 5: Relationship between NDVI and malaria anomalies. The coefficient of determination (R^2) values indicates the proportion of variability in the malaria anomalies that is accounted for by the NDVI.

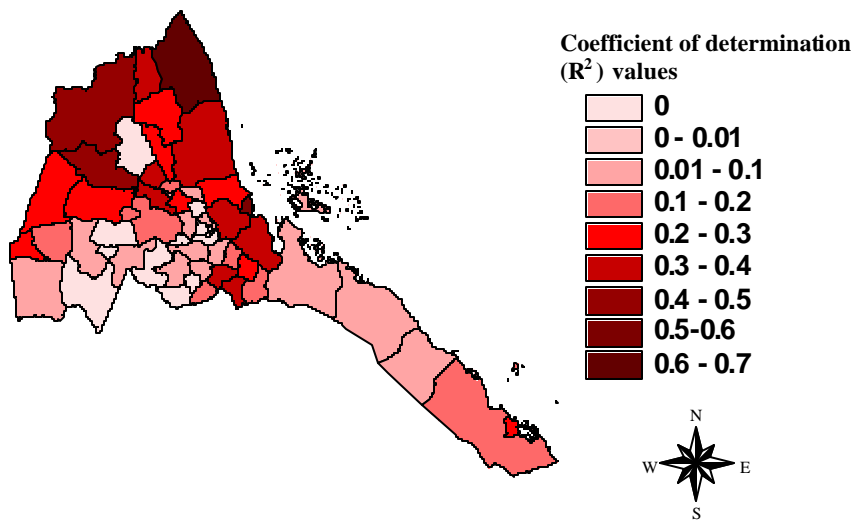


Table 1: Types and sources of observed or forecast climate data for Eritrea

Climate data	Coverage	Spatial resolution	Temporal resolution	Time period	Source
Rainfall gauge data	23 meteorological stations	Point data	Monthly cumulative rainfall	Jan 1992 to Dec 2003	Meteorology Office of the Civil Aviation Department.
Climate Prediction Center Merged Analysis of Precipitation (CMAP) version 0407	Complete coverage	2.5 ×2.5 degree grid	mm per day for each month.	1979 to present	(http://www.cpc.ncep.noaa.gov/products/global_precip/html/wpage.cmap.html).
Normalized Difference Vegetation Index (NDVI) version “e” product	Complete coverage	~8km	Maximum value of three dekads in one month	Jul 1981 to present	Provided by USGS ADDS and made available via the IRI data library at http://iridl.ldeo.columbia.edu/SOURCES/USGS/ADDS/NDVI/NDVIe/.dekadal/.maximum/.NDVI/
Interpolated rainfall gauge data	Complete coverage	~5km	Monthly cumulative rainfall	1959-1998	Climate Research Unit of the University of East Anglia precipitation dataset. ³⁷
SSTs	Global oceans	2 X 2 degree grid	Monthly mean	1979 to present	NOAA ERSSTv2 available via IRI data library at: http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCDC/.ERSST/.version2/

ECHAM 4.5	Global	2.8 X 2.8 degree grid	Monthly mean	1979 to present	Max Planck Institute, Hamburg Germany, model run at IRI
-----------	--------	-----------------------	--------------	-----------------	---

Table 2: Number of health facilities per cluster and number of reports received for each year

	Five clusters corresponding to Figure 3				
	1 Very low incidence year around (Oct peak)	2 Low incidence year around (Oct peak)	3 Moderate incidence West (Oct peak)	4 High incidence (Oct peak)	5 Moderate incidence East (Jan peak)
Number of Health Facilities	42	103	38	24	35
Number of Reports 1996	339	710	140	222	102
Number of Reports 1997	373	942	268	248	236
Number of Reports 1998	395	1012	334	239	289
Number of Reports 1999	406	1089	352	244	329
Number of Reports 2000	419	1037	284	219	288
Number of Reports 2001	437	1106	335	260	326
Number of Reports 2002	417	1095	400	247	349
Number of Reports 2003	439	1108	405	265	381

Table 3: Average of reports per health facilities per cluster and per year.

	Five Clusters corresponding to Figure 3				
	1 Very low incidence year around (Oct peak)	2 Low incidence year around (Oct peak)	3 Moderate incidence West (Oct peak)	4 High incidence (Oct peak)	5 Moderate incidence East (Jan peak)
Number of Reports per Facility 1996	8.07	6.89	3.68	9.25	2.91
Number of Reports per Facility 1997	8.88	9.15	7.05	10.33	6.74
Number of Reports per Facility 1998	9.40	9.83	8.79	9.96	8.26
Number of Reports per Facility 1999	9.67	10.57	9.26	10.17	9.40
Number of Reports per	9.98	10.07	7.47	9.13	8.23

Facility 2000					
Number of Reports per Facility 2001	10.40	10.74	8.82	10.83	9.31
Number of Reports per Facility 2002	9.93	10.63	10.53	10.29	9.97
Number of Reports per Facility 2003	10.45	10.76	10.66	11.04	10.89