Malaria and Early African Development: Evidence from the Sickle Cell Trait

Emilio Depetris-Chauvin
Brown University

David N. Weil
Brown University and NBER

October 2011

Abstract
We examine the effect of malaria on economic development in Africa over the very long run. Using data on the prevalence of the mutation that causes sickle cell disease we measure the impact of malaria on mortality in Africa prior to the period in which formal data were collected. Our estimate is that in the more afflicted regions, malaria lowered the probability of surviving to adulthood by about ten percentage points, which is roughly twice the current burden of the disease. The reduction in malaria mortality has been roughly equal to the reduction in other causes of mortality. We then ask whether the estimated burden of malaria had an effect on economic development in the period before European contact. Examining both mortality and morbidity, we do not find evidence that the impact of malaria would have been very significant. These model-based findings are corroborated by a more statistically-based approach, which shows little evidence of a relationship between malaria ecology and population density or other measures of development, using data measured at the level of countries or ethnic groups.

We are grateful to Ronald D. Lee, Andrew Mason, Gordon McCord, and Nathan Nunn for sharing data, to Fatima Aqeel, Federico Droller, and Scott Weiner for research assistance, and to Quamrul Ashraf and seminar participants at Harvard Center for Population and Development Studies, Indiana University, and Pennsylvania State University for helpful comments.
It is impossible to understand the pattern of comparative economic development in the world today without understanding comparative development in the past. Consider, for example, a horizon of 500 years. Countries and regions that were highly developed as of the year 1500 are, for the most part, among the most developed today. Exceptions to this regularity, such as China, tend to be growing quickly. Taking into account flows of population over the last half millennium makes this correlation even stronger: Countries populated by people whose ancestors lived in the most developed countries are most likely to rich today. Looking within countries, people who are descended from parts of the world that were highly developed in the year 1500 are on average higher up in the income distribution than people descended from regions that were not developed.

Going back further back in time, there is still strong predictive power of past development for present outcomes. Comin, Easterly, and Gong (2010) show that not only is a country’s level of technology from 500 years ago predictive of income today, but so is the level of technology 2,000 or 3,000 years ago. Hibbs and Olsson (2004) show that the date at which the transition from hunting and gathering to settled agriculture took place is predictive of a country’s income today.

The fact that development in the past is so predictive of development today suggests two possible theories. First, it may be that the same factors which held back development in previous historical eras are still operative in the present. Examples of such factors are climate associated with certain geographical locations or slowly changing aspects of culture or institutions that might be associated with different population groups. Alternatively, it may be that the specific factors that caused past underdevelopment are no longer relevant today, but that the fact of past underdevelopment itself is causal of current underdevelopment. For example, it could be that the early development advantage of the Eurasian land mass arose from the historical presence of plentiful species of large seeded grasses and domesticable animals, as argued by Diamond (1997), but that the continuation of the development gap between Eurasia and other regions results from the effect of colonial institutions that Europeans were able to impose on much of the rest of the world as a result of this initial advantage.

Whichever of these theories is correct (and obviously it is possible for both of them to have some validity), there is clearly much to be learned by looking at the roots of development differences in the past. In this paper, we examine the historical impact on development of malaria. Malaria is one of the most significant diseases in the world today in terms of its humanitarian impact. With some 225 million annual cases, it accounts for one in five child deaths in Africa. Malaria’s control and impacts are widely studied by biologists and social scientists. Economists actively debate its role in affecting growth in the modern world. However, as the above discussion makes clear, it would be possible that even if malaria were not important in affecting economic development today directly, it could nonetheless have been an important determinant of growth historically.

2 The widely quoted estimate of Gallup and Sachs (2001) is that malaria reduces growth of GDP per capita by 1.3% per year in the African countries most afflicted. See Weil (2010) for an extensive critique of this literature.
Analysis of the role played by malaria in shaping historical development is severely hampered by a lack of data. Biologists only came to understand the nature of the disease in the late nineteenth century, and even today, trained medical personnel have trouble distinguishing between malaria and other diseases without the use of microscopy or diagnostic tests. Accounts from travelers and other historical records provide some evidence of the impact of the disease going back millennia, but these are hardly sufficient to draw firm conclusions. As discussed below, there do exist data (“malaria ecology”) which measure the extent to which the environment in different geographical regions is supportive of malaria. One can look at the empirical relationship between this malaria ecology measure and economic development, either currently or historically. We discuss such a statistical approach below. However, such an approach faces severe limitations. For example, it is difficult to know whether one has controlled for correlates of malaria ecology that might independently influence the process of development. These correlates could be other diseases or factors such as tropical climate, which might affect agriculture, or the effect of tropical climate on European mortality, and thus current institutions, as argued by Acemoglu, Johhson, and Robinson (2001).

In this paper we address the lack of information on malaria’s impact historically by using genetic data. In the worst afflicted areas, malaria left an imprint on the human genome. Specifically, we look at the prevalence of the gene that causes sickle cell disease. Carrying one copy of this gene provided individuals with a significant level of protection against malaria, but people who carried two copies of the gene died before reaching reproductive age. Thus the degree of selective pressure exerted by malaria determined the equilibrium prevalence of the gene in the population. By measuring the prevalence of the gene in modern population, we can thus back out an estimate of the severity of malaria historically. We compare these estimates to the burden of malaria today and discuss how malaria compares to other determinants of premature mortality, both in Africa and elsewhere.

With estimates of the extent of malaria mortality in hand, we then turn to look at the impact of the disease on economic development. Because of the potential correlation of malaria with other factors affecting development, we eschew a regression-based approach. Instead, we follow the path of using the machinery of standard economic modeling. Ashraf, Lester, and Weil (2007) for example, apply this technique to look at the effect of disease on economic growth in the world today. In the current paper, we focus in particular on quantifying the economic effects of malaria mortality, although we also address the effect of morbidity resulting from the disease.

In studying the role of malaria in long run development, we are also inevitably studying the long run development of Africa. Both historically and today, Africa has been the major focus of the disease. Indeed, malaria was not present in the tropical regions of the new world until it was accidentally brought there by Europeans (McNeill, 1977). Historians of Africa attribute a large role to diseases in general, and malaria in particular, in shaping development (Akyeampong, 2006). For example, Webb (2006) describes malaria, along with trypanosomiasis (sleeping sickness) as having profoundly influenced African patterns of settlement as well as culture. The

---

genetic data that we use gives us a unique opportunity to assess the role of malaria in shaping geographic heterogeneity in Africa.4

The rest of this paper is organized as follows. In Section 1, we present a statistical analysis of the relationship between malaria ecology and a variety of measures of development, both contemporary and historical. Of particular interest, in this section we explore the use of data at the level of ethnic groups as an alternative to more common country-level analyses. In section 2 we discuss the biology of the link between malaria and sickle cell disease, and we examine the empirical relationship between measures of the two. Section 3 presents and applies our model for using the current level of sickle cell prevalence to estimate the historical burden of malaria, including comparisons of malaria’s historical burden to its current level and to the historical burden of other diseases. In section 4 we turn to the question of how malaria effected development. We examine a model in which the economic cost of the malaria mortality can be evaluated, and also discuss the effect of malarial morbidity. Section 5 concludes.


Malaria is caused by the plasmodium parasite, which is transmitted to humans through the bite of a female anopheles mosquito. Early symptoms of malaria include fever, chills, severe headache, and vomiting. In severe cases these are followed by respiratory distress, severe anemia, or cerebral malaria. Infants are protected from the disease in the first few months of life by a combination of material antibodies and characteristics of the structure of fetal hemoglobin. In malaria endemic areas, most children have developed substantial immunity by the age of five.

Africa currently accounts for 85 percent of world malaria cases and 90 percent of world malaria deaths. The geographical pattern of malaria's severity is largely determined by the climactic conditions that support mosquito breeding as well as by the mix of mosquito species present. There are significant differences in the vectorial capacity among the approximately 20 species of anopheles that transmit malaria, based on factors such as the mosquito's preferred targets, biting frequency, and lifespan. The most effective vector, Anopheles Gambiae, is the principal vector in Africa.

\[4\] Weil (2011) paints a picture of African development in 1500, both relative to the rest of the world and heterogeneity within the continent itself, using as his indicators population density, urbanization, technological advancement, and political development. Ignoring North Africa, which was generally part of the Mediterranean world, the highest levels of development by many indicators are found in Ethiopia and in the broad swathe of West African countries running from Cameroon and Nigeria eastward along the coast and the Niger river. In this latter region, the available measures show a level of development just below or sometimes equal to that in the belt of Eurasia running from Japan and China, through South Asia and the Middle East, into Europe. Depending on the index used, West Africa was above or below the level of development in the Northern Andes and Mexico. Much of the rest of Africa was at a significantly lower level of development, although still more advanced than the bulk of the Americas or Australia.
Kisezewski et al. (2004) construct an index of “malaria ecology” that takes into account both climactic factors and the dominant vector species to give an overall measure of how congenial the environment is to the spread of malaria. The index does not take into account public health interventions such as swamp-draining or differences among countries in economic development or health care infrastructure. In other words, it represents the component of malaria variation among countries that is exogenous to human intervention. The index is calculated for grid squares of one half degree longitude by one half degree latitude. Figure 1 show data for the whole world, and Figure 2 shows a close-up of Africa. With the exception of New Guinea and some areas of southeast Asian, Africa is the only part of the world in which the index reaches its highest levels. Areas in which malaria played a significant role historically but has now been eradicated, such as Greece, Southern Italy, and the American South, are all seen to have relatively low values for the malaria ecology index. Within Africa, there is substantial variation in the index. Figure 3 depicts the partial regression line for the relationship between malaria prevalence in 2008, measured as notified cases of malaria per 100,000 population, and average malaria ecology at the country level, controlling for continent dummies. The implied coefficient estimate shows the strong, positive and statistically significant relationship between the two measures.

1.1. Country Level Analysis

Table 1 present the OLS coefficient estimates for different specifications of the following equation:

\[
\ln(GDP_{pc})_i = \alpha + \beta M_i + \gamma M_i \times OldWorld_i + \theta OldWorld_i + \tau G_i + \pi K_i + \varepsilon_i \quad (1)
\]

Where the subscript \(i\) denotes country; GDPpc is real GDP per capita in 2000; \(M\) is the mean of the malaria ecology index; OldWorld is a dummy variable that takes the value 1 if the country is not in the Americas or Oceania; \(G\) is a vector of geographic controls including distance to the nearest coastline, absolute latitude, percentage of the land surface area of the country that has any of the four Köppen-Geiger tropical climates, and mean elevation; \(K\) is vector of continent dummies (to avoid perfect multicollinearity only a dummy for the Americas is included if the specification states \(\theta \neq 0\)). \(\varepsilon\) is the error term that is allowed to be heteroskedastic (we estimate robust standard errors in all the specifications).

Results in column 1 include only geographical controls and suggest that, ceteris paribus, 1-point increase in the malaria ecology index is associated with a 5.7 percent decrease in GDP per capita. The addition of continent dummies in column 2 reduces the size of the coefficient by roughly half without affecting its statistical significance. When we exclusively focus on this statistical relationship within Africa in column 3, we find a coefficient of similar magnitude but weaker in a statistical sense (only significant at the 10 percent level). This result is striking since we expect the negative effect of malaria on development to be stronger in Africa. Nonetheless, the sharp drop in sample size in this specification is likely to partially account for this reduction of statistical significance.

In column 4 we focus on the differential effect of malaria in the Old World (defined as all the countries with the exception of the Americas and Oceania). The coefficient estimate for the
interaction term malaria ecology times Old World dummy is not statistically different from zero suggesting that malaria has no differential effect on real GDP per capita in the Old World (relative to the Americas and Oceania). Moreover, we cannot reject the null of hypothesis (p-value 0.51) that the sum of the coefficients for malaria ecology and its interaction is statistically equal to -0.033 (the coefficient obtained for malaria ecology in the specification of column 2). The last specification in column 5 restricts the sample to the Old World. According to the coefficient estimate, a 1-point increase in the malaria ecology is associated with 3.7 percent increase in real GDP per capita; a similar magnitude to the one found in column 2 with the unrestricted sample. Overall, malaria ecology is predictive of low income today. Interestingly, the statistical relationship between malaria ecology and GDP per capita does not seem to be different when we distinguish between New World and Old World. Moreover, the implied semi-elasticity of real GDP per capita with respect to malaria ecology does not differ statistically when we take into account the full sample of countries or we solely focus on African countries.

Now we focus entirely in the Sub-Saharan Africa to examine the association between malaria and population density, rather than income, for the usual Malthusian reasons. When explaining variation across country, it is important to account for other factors affecting population density, such as differences in geography and land quality, and have nothing to do with different level of technologies in a Malthusian world. While we do include geographic controls we pay special attention to difference in land quality. Based on high-resolution datasets for croplands, climate and soil characteristics, Ramankutty et al. (2002) develop an index of land suitability for agriculture. This index represents the probability that a particular grid cell (of the size of about 50 x 50 km) may be cultivated. The authors assume that suitability for cultivation is solely a function of climate (temperature, precipitation, and potential sunshine hours) and soil properties (total organic components measured by carbon density and nutrient availability based on soil pH). Other biophysical and socio-economic factors, such as topography and irrigation or market price and incentive structure, are omitted. It is important to note that the index does not account for the productivity of a particular piece of land, but simply whether the characteristics of the land are favorable for crop cultivation. For the rest of the paper we refer to this variable as land quality.

In Table 2 we present cross-country regression outcomes for different specifications of the following equation:

\[
\ln(P)_i = \alpha + \beta M_i + \gamma L_i + \delta M_i \times L_i + \tau G_i + \pi' I_i + \epsilon_i
\]

(2)

Where the subscript \(i\) denotes country; \(P\) is population density; \(M\) and \(G\) are the same variables as defined for equation (1); \(L\) is the mean value of land quality index; \(I\) is a vector of dummy variables to account for the colonial power ruling the country; \(\epsilon\) is the error term that is allowed to be heteroskedastic. Specifications in columns 1 to 3 use population density data for 1500 CE as the dependent variable so \(\pi' = 0.\) The point estimate in column 1 suggests that 1 point increase in malaria ecology is statistically associated with 5.8 percent increase in population density in 1500 CE. However, when we add land quality and its interaction with malaria ecology

5 The interaction term is based on demeaned variables so the main coefficients for malaria are directly comparables after the addition of the interaction term.
6 We use population density data in 1500 CE from Chanda and Putterman (2007).
none of the coefficients are statistically significant at standard levels. Adding the geographic controls, in column 3, does not change the statistical insignificance of all the coefficients. In sum, country level data for Sub-Saharan Africa does not present evidence of any effect of malaria ecology on the commonly used measure of development in 1500 CE.

In columns (4) to (7) we focus on more modern data. Our dependent variable is the log of population density in the year 1950 from Maddison (2003) which is the furthest back in the past we could get modern data on population density at the country level for Sub-Saharan Africa. Coefficient estimates in columns 4 to 6 show no statistical relationship between malaria ecology (either separately or interacted with land quality) and population density in 1950 regardless of the addition of geographic controls. However, when we control for the colonial rule of the country, in column 7, we find a statistical significance and positive association between malaria ecology and the log of population density (at the 10 percent level). It is important to note, however, that this apparent lack of statistical significance of malaria ecology could be driven by measurement error on population density data. But most importantly, the aggregation of malaria ecology data at the country level might be masking heterogeneity within country boundaries. Especially for large countries, such as Nigeria, Mali or Congo DR, an average measure of the malaria ecology index is probably unable to capture extreme values of the spectrum of the index and their effect on population density or any measure of development. Moreover, modern country borders are probably not very meaningful not only for 1500 CE but even for the beginning of the 20th century. Finally, population density varies widely within modern country borders. In next section we exploit cross-ethnic group variation in population density, malaria ecology, and a rich set of covariates to shed some light on the effect of malaria on economic development within Africa.

1.2. Ethnic Group Analysis

Using population figures from Murdock (1967)’s *Ethnographic Atlas* (EA) we construct population density figures at the ethnic group level for the colonial period. The EA is a database on 1167 ethnic groups of six different regions of the world, including Sub-Saharan Africa. This database includes a variety of information from the level of subsistence economy to the degree of political integration of each group. For some of the African ethnicities, the EA also includes information on population size and year of its estimation. We identified figures and date for over 250 African ethnic groups. In order to compute the population density we use total land area for each ethnic group as implied from shape-file from the digitalized version of the ethnicities’ Map in Murdock (1959). Since there is no perfect match between ethnicity names in Murdock (1967) and Murdock (1959)’s map we were able to locate only 206 ethnicities with population figures. The date of estimation varies in the range period 1880-1960. Figure 4 presents the histogram for the year of population figures collection. Almost half of the sample belongs to the first half of the 20th century and earlier whereas the other half is clustered in the 1950s.

Figure 5 displays the population density for the 206 ethnicities in our sample. We divide the sample in quintiles of the number of people per square kilometer at the ethnicity level. Ethnic

---

7 Europeans would have not taken into account local conditions and ethnic characteristics when drawing current national borders in Africa, which explains the existence of several ethnic groups who were partitioned into different countries (Michalopoulos and Papaioannou, 2011).
groups from Central, East and West Africa are more prevalent in our sample whereas only 5 groups belong to North Africa. Except for the cluster of highly densely populated groups near Lake Victoria, the quintiles seem to be evenly distributed across in the Sub-Saharan Africa. Figure 6 shows a close-up of West Africa and from its inspection we can note that there is variation in population density even within that region.

1.2.1. The effect of Malaria on Population Density during Colonial Time

In order to assess the effect of malaria environment in our measure of ethnicity prosperity during the colonial time we estimate several different specifications of the following equation:

\[
\ln(P)_i = \alpha + \beta M_i + \gamma L_i + \delta M_i \times L_i + \theta' T_i + \rho' C_i + \tau' G_i + \pi' I_i + \varepsilon_i
\]  

(3)

Where the subscript \(i\) denotes ethnic group; \(P\), \(M\), and \(L\) were defined above; \(T\) is a vector of 8 decade dummies for the period 1880 -1950; \(C\) is a vector of 10 climate dummies; \(G\) is a vector of geographic controls including geodesic distance of the centroid of the historical homeland of each ethnic group from the nearest coastline, mean elevation of terrain, and absolute latitude; \(I\) is a vector of external influence including dummy variables for the colonial power ruling the land of each ethnic group, and two variables accounting for the importance of the Atlantic and Indian Ocean slave trades for the period 1400-1900; \(\varepsilon\) is the error term that is allowed to be heteroskedastic. Note that \(T \epsilon \) [1880, 1950] and \(T_i\) takes value 1 if the year for the population figure for \(i\) belongs to the decade \(T\). The climate dummies embodied in \(C\) follow the Koeppen-Geiger classification.

Table 3 presents the first statistical results. In column 1 we impose \(\gamma = \delta = \theta' = \rho' = \tau' = \pi' = 0\) and the coefficient estimate for \(\beta\) suggests a positive and statistically significant correlation between malaria ecology and population density. Although only significant at the 10 percent level, the point estimate from our semi-log specification indicates that one point increase in scale of the malaria ecology index would increase population density in around 2.5 percent. Therefore, one standard deviation increase in the malaria ecology (i.e: 9.4 in our sample of ethnic groups) imply a 25 percent increase in population density. As expected this statistical relationship is not affected by the inclusion of decade dummies in column 2. Interestingly, the introduction of climate fixed only marginally affects the standard errors and the size of the coefficient, improving the t-ratio in column 3 though. Taking all together malaria ecology, the decade dummies and the climate fixed effect explain 35 percent of the variation of the log of population density.

High values of the malaria ecology index coincide with high values of the land quality index in some part of Africa, such as the Sahel region. This correlation might be partially accounting for a common factor, that is access to water, that also correlates with population density. We add a measure of land quality at the ethnic group level in column 4. The positive and significant correlation between malaria ecology and population density is not statistically affected by the addition of this new control. An increase of one standard deviation in mean land quality (i.e:
0.24 points in the scale of land quality index) is associated with an increase of 32 percent in population density. 

Now we turn to the relationship between land quality and malaria and their common effect on population density. We add the interaction term mean malaria ecology – mean land quality in column 5. The relationship between malaria ecology and population density becomes statistically stronger since now the coefficient estimate is statistically significant at the 1 percent level with the addition of the interaction term. Note that in order to make comparable the size of the coefficient for the malaria semi-elasticities in column 5 and 4 we need to demean the explanatory variables. Taking the mean value of land quality (i.e: 0.47), the coefficient estimates in column 5 suggest that 1-point increase in the malaria ecology would positively affect population density by 2 percent (from doing 0.0919 – 0.154*0.47). Nevertheless, there is a threshold in the value of land quality at which the effect of malaria in population density becomes negative (i.e: mean land quality = 0.6). More than 60 ethnic groups (approximately 1/3 of our sample) present a mean land quality index above this threshold. Half of these ethnic groups are located in East Africa and 1/4 in West Africa.

Even though our measures of malaria environment and land quality can be understood as measures not affected by human intervention so we can rule out an endogeneity problem due to reverse causation from population density to our two main explanatory variables, we can still have some omitted variable bias. For instance, the statistical relationship between geographic factors and development, either through their direct or indirect causal mechanisms, is well documented in the empirical literature. We start addressing this issue by adding the geodesic distance of the centroid of the historical homeland of each ethnic group from the nearest coastline, mean elevation of terrain, and absolute latitude as geographic controls in column 6. The introduction of the 3 controls reduces the size of the main land quality effect and its interaction with malaria. However, both coefficients remain statistically significant at the 1 percent level. The coefficient estimate for malaria ecology is not statistically affected. The threshold of land quality to make the effect of malaria negative is now 0.8 and only 10 percent of the sample presents values of land quality above that threshold.

Our population data belongs to the African colonial time. If some ethnic groups located in places where the disease environment and other geographic factors handicapped development were also the most affected by the colonial rule, then no taking into account cross-ethnicity variation in colonial power might lead to an important bias in the estimates of the effect of malaria and its interaction with land quality. Nonetheless, the addition of dummies variables to account for the

---

8 Note that this result is somehow consistent with previous finding on the relationship between land productivity and population density during a Malthusian regime. Using data for 1500 CE at the country level Ashraf and Galor (2010) estimate the elasticity land productivity – population density to be approximately 0.5. This means that a 1 percent increase in land productivity would lead to an increase of 0.5 percent in population density. Assuming that our land quality index and their land productivity variable are measuring the same, we can calculate how much population density would increase if we increase land quality in a percentage amount consistent with one standard deviation increase over the mean. The mean of land quality index in our sample is 0.47 and its standard deviation is 0.24, thus, increasing one standard deviation implies a 50 percent increase over the mean of land quality. According to Ashraf and Galor (2010), 50 percent increase in land productivity implies around 25 percent increase population density which is not substantially different from our 32 percent increase using data at the ethnic group level in Africa.
effect of the colonial power nationality and slave trade in columns (7) and (8), respectively, does not affect the previous results. In the last specification, we omit from the sample 5 ethnic groups that belongs to North Africa in column (9). The results are similar to and consistent with the ones obtained in previous specifications.9

1.2.2. Other Measures of Prosperity

We turn now to the relationship between malaria on other measures of ethnic prosperity. We exploit a rich source of ethnic information provided by the EA (1967) and estimate a new version of equation (3) where our dependent variable is now a given measure of ethnic prosperity. Table 4 presents the main results. All the specifications include decades dummies, climate fixed-effect, geographic and colonial rule controls. Since we do not longer rely on the availability of population figures, the sample size for each specification will depend only on the availability of the ethnographic variable for ethnic groups that can be located in the Murdock’s map. Note that now the assignment of the decade dummy is based on the year to which the ethnographic data pertain (Murdock, 1967). Figure 7 presents a histogram showing that one fourth of the data belongs to the period 1820-1900, 65 percent to the first half of the 20th century, and only 10 percent (39 observations) to the 1950s. Note, however, that adding a time control in our regressions is probably not crucial for most of these measures of ethnic traits since the political and economic structure of a given ethnic group was likely to be determined in earlier periods (i.e: pre-colonial times) and transmitted inter-generationally.

The dependent variables in columns 1 to 5 in Table 4 are categorical variables (see appendix for exact definitions) listed in order of increasing prosperity levels. The dependent variable in the last specification is the first principal component of the 5 previous measures. We first study the statistical relationship between malaria, its interaction with land quality and the mean size of local communities for each ethnic group. Column 1 in Table 4 suggests that malaria ecology has a direct positive effect on the mean size of the communities but its interaction effect with land quality is not statistically significant (although its signs is consistent with previous findings). The point estimate for the main coefficient in column 2 suggests no statistical relationship between malaria and the settlement pattern of the ethnic group (i.e: whether the ethnic groups were fully nomadic, semi-sedentary or lived in complex settlement). Malaria ecology is negatively and statistically related to this measure of prosperity, however, when interacted with the average quality of the land inhabited by the ethnic group.

In column 3 we relate malaria ecology with a measure of sophistication of the political institutions at the ethnic group level. The statistical results suggest that ethnic groups located in areas with worst malaria ecology tend to have most sophisticated political institutions beyond the local communities (i.e: large states in the highest possible category in the spectrum of sophistication). In line with previous findings in Table 3, malaria ecology interacted with land quality presents a negative and strongly statistically significant coefficient suggesting that malaria would have a negative effect for ethnic group located in the most fertile areas. We find

9 We also experimented with other specifications including region fixed dummies (i.e: West, East, Central, North, and South) and ethnographic region dummies (i.e: African Hunters, South African Bantu, Central Bantu, Northeast Bantu, Equatorial Bantu, Guinea Coast, Western Sudan, and Nigeria Plateau). The coefficient estimates are not substantially affected.
similar results when taking into account the relationship between malaria ecology and the level of subsistence of the economy (i.e: from mostly reliance on gathering to mostly agricultural economy) in column 4. Interestingly, malaria ecology is not statistically associated to the level of intensity of agriculture (column 5). Finally, using the first principal component of the previous measures of prosperity as a dependant variable in column 6 suggests that malaria ecology is only statistically related to the prosperity of the ethnic group through its interaction effect with land quality (significant at the 10 percent level).

2. Malaria and Sickle Cell Disease

Several mutations have arisen in human populations that provide resistance to malaria. These include the mutation causing thalassemia, which is present in Mediterranean, Arab, and Asian populations; the absence of the Duffy blood group in west Africa; hemoglobin E in Southeast Asia; and hemoglobin C in West Africa (Allison, 2002; Nelson and Williams, 2006). The most important such mutation is the one that causes sickle cell disease.

Sickle cell trait is a mutation in the gene that produces hemoglobin, the oxygen-carrying component in red blood cells. Individuals carry two copies of this gene, one received from each parent. Individuals who carry one normal copy of the gene (referred to as A type) and one copy with the sickle cell mutation (S type) are carriers of the disease. In individuals of the AS phenotype, a fraction of the hemoglobin in their red blood cells have an abnormal structure. In individuals who have two copies of the sickle cell gene (SS phenotype), almost all hemoglobin molecules are of the abnormal type.

In conditions of inadequate oxygen supply (hypoxia), hemoglobin produced by the S gene becomes rigid, leading to a characteristic sickle shape of red blood cells. Carriers of sickle cell trait generally do not suffer many adverse effects. However, there can be negative consequences from sickling in conditions of low oxygen such as unpressurized airplane flights and extremely rigorous exercise (Motulsky, 1964). In individuals of the SS phenotype, such sickling of red blood cells is far more common, leading to acute episodes of disease in which abnormally shaped cells restrict blood flow to organs, severe damage to the spleen, and anemia. In 1994, life expectancy for SS children in the United States was 42 years for males and 48 years for females. In the absence of modern medical care, individuals of the SS phenotype are not able to survive to adulthood.

The sickle cell mutation is relevant to malaria because infection of a red blood cell with the malaria parasite leads to hypoxia. In individuals of the AS phenotype such blood cells sickle and are then eliminated by the body's immune system, lessening the burden of infection. Carriers of the sickle cell trait are particularly resistant to severe malarial episodes; they are less resistant to mild cases. The mechanism by which AS carriers are protected from malaria is different than the acquired immunity that both AA and AS individuals achieve following repeated exposure to the disease.

---

The benefit that possessing a single copy of the sickle cell gene conveys counterbalances the biological cost incurred when homozygous SS children are stricken with sickle cell disease. An individual of the AS phenotype is more likely to reach adulthood than is an individual of the AA phenotype, but the former is also more likely to see his/her child die of sickle cell disease. This is known as a heterozygote advantage or balanced polymorphism. As shown more formally below, the stronger is the pressure of malaria on survival, the more advantaged are individuals who carry the S gene, and in equilibrium, the higher the percentage of the population who will be carriers. Indeed, it was the correlation of high prevalence of the sickle cell gene and the presence of malaria that first led scientists to understand the protective role of the sickle cell mutation.

As will be seen in the next section, the underlying genetic mechanism by which the sickle cell trait is transmitted provides a means of mapping sickle cell prevalence into an estimate of the mortality burden of malaria. To map this burden in turn into an estimate of the relevance of malaria ecology, we study the relationship between sickle cell prevalence and malaria ecology. In order to do this, we have constructed a new dataset on sickle cell prevalence.

**Data**

We georeferenced 670 population surveys from Livingstone’s (1987) extensive compilation on sickle cell frequencies. Although the original data includes different studies that used different techniques, Livingstone states that all results for the sickle-cell frequencies should be comparable. Livingstone original dataset includes information on the name and location of the population tested, the number of individuals with sickle cell trait, AS and S hemoglobin type, and the total number of individuals tested. We considered as AS the people classified as S hemoglobin type. Our georeferencing process involved the assignment of a geographic coordinate to each population survey in Livingstone’s original dataset. In many cases, the same population or location was surveyed multiple times, leading in some cases to different figures of the prevalence of the genetic mutation. Since there are no indisputable criteria to choose the sample that is the most representative of a given location, we included all the observations in our dataset.

Figure 8 presents the geographic distribution of our sickle cell prevalence data on the malaria ecology map of Africa. Blue dots represent locations where sickle cell prevalence is below 10 percent. Almost 40 percent of our sample presents sickle cell prevalence in that range and seems to be evenly dispersed throughout of Africa. Above 20 percent of our sample presents prevalence in the range 10 – 20 percent. Although there are two cases in North Africa, this group appears to be more prevalent in the Sub-Saharan region and Madagascar. The map also suggests that high levels of sickle cell trait do not occur in the southern part of the continent. The west part of Africa, from Cameroon to Senegal, shows a large heterogeneity in the prevalence of the sickle cell trait, suggesting both almost absence of the mutation in some places and very high values of prevalence (above 40 percent) in others. A large contiguous area with values between 20 to 50 percent prevalence (in orange, yellow, and red dots) is observed in Central Africa from Congo Republic to Zambia. High values of the sickle cell trait are also present in the proximity of Lake Victoria. Finally, two outliers are denoted with red triangles for which prevalence is above 80 percent (one is located in the Negueah Dubreka district in Guinea, and the other in Uganda). Both cases present small sample sizes (based on the number of people tested).
Some pitfalls are presented in our dataset based on Livingstone’s work. In many cases the survey belonged to people in hospitals. The inclusion of non-random population samples might bias sickle cell frequencies. Nonetheless, according to Livingstone, an examination of the frequencies for hospital patient and other individual in the same population, where these were available, did not seem to indicate as much discrepancy or difference as might have been expected. In addition, we could not discriminate between indigenous and non-indigenous population. This could be problematic for the final goal of assessing the historical impact of malaria using this genetic data if non-indigenous people, with a given equilibrium level of sickle cell prevalence, takes a migration decision based on the characteristics of the malaria environment of either the destination or the origin. Moreover, Livingstone (1958) suggests that the distribution of sickle cell trait in some part of West Africa might be the result of not only a selection mechanism but also gene flow due to migration and mixture.

After the completion of the georeferencing process of our sample we became aware of the existence of a more detailed S gen frequency dataset that would help to overcome the aforementioned pitfalls. Piel et al (2010) present a global geodatabase of S gene frequency based on comprehensive electronic search of academic publications presenting S gene frequency figures. Each reference finally included in the dataset met the first criteria that the population surveyed were representative of the indigenous population of a particular location. To avoid other source of bias, they also state that the dataset excludes non-representative populations, such as samples of sickle cell patient. Piel et al (2010) assigned a geographic coordinate to all samples with the distribution of AS and AA genotypes that met the strict inclusion criteria. Using a Bayesian model-based geostatistical framework they then create a continuous map of the sickle cell gen frequency resulting in 10 km by 10 km resolution global raster grid. Figure 9 presents the map. It is important to note that throughout our paper we use the terminology sickle cell trait or prevalence (that’s the fraction of people who are carriers of the allele S) instead of sickle cell gen (or S allele) frequency as in Piel et al (2010). Since very few S homozygotes survive to adulthood, the sickle cell gen frequency would be close to 1/2 of the sickle cell trait frequency. Even though their maximum values for the implied prevalence of the genetic mutation do not reach 40 percent (roughly 20 percent for the S gene frequency), whereas we have figures well about that number in our dataset, only 2 percent of our sample is above that threshold. Consistent with our map presented in Figure 8, the maximum levels of sickle cell prevalence are located in Central Africa. West Africa also presents an important heterogeneity whereas medium level frequencies are also located in the proximity of Lake Victoria.

Although the data from Piel et al (2010) seems to be the best current data for academic use, is not publicly available yet. Therefore, we use our dataset to study the empirical relationship between sickle cell prevalence and malaria ecology. Hopefully, we expect to have access to the Piel et al’s data very soon so the following empirical work should be taken as very preliminary and subject to potential problems due to the quality of the data. Note that the availability of the grid square map would allow us to study this relationship in different levels of aggregation (i.e: country, ethnic groups, and even 10 by 10km square grid).

Relationship Between Sickle Cell Prevalence and Malaria Ecology
In order to study the statistical relationship between sickle cell prevalence and malaria ecology we matched each georeferenced (i.e: latitude and longitude) prevalence figure to its malaria ecology counterpart. Using ArcGis we intersected the latitude-longitude coordinate of each data point in our dataset with raster data of 0.5 degree by 0.5 degree resolution (approximately 50 km by 50 km) for malaria ecology. In other words, we located our sickle cell data points in their corresponding 0.5 degree by 0.5 degree grid square and assign its value of the ecology index. Table 5 presents the first piece of evidence regarding the statistical relationship between sickle cell prevalence and malaria ecology. Point estimate from column 1 suggest that one standard deviation increase in the malaria ecology index (8.4 points) is statistically associated to 0.83-points increase in the sickle cell prevalence (significant at the 10 percent level). When we weight the regression by the number of people tested in each sample in column 2, we find a larger and statistically stronger association that shows that one standard deviation increase in the malaria ecology index is associated with a 4.3 point increase in the sickle cell prevalence (significant at the 1 percent level). Note that, according to this specification, Malaria ecology explains 20 percent of the variation of the prevalence of the genetic mutation. A close inspection of the scatter plot in Figure 10 suggests a non-linear relationship between the two measures. A weighted (the larger the size of the dot, the larger the population sample) quadratic function on malaria ecology seems to be a better fit relative to its linear counterpart.\footnote{Because we do not know whether this inverted-U shaped relationship will hold up using the improved data from Piel et al., we hold off on trying to interpret its meaning.}

3. Measuring the Historical Burden of Malaria Using Data on the Sickle Cell Trait

3.1 Simple Model

Our goal is to examine what the prevalence of the sickle cell trait among African populations tells us about the impact of malaria historically. As described above, every adult carries two alleles. These can be either sickle cell (S) or normal (A).\footnote{We ignore other mutations.} A person with the SS phenotype will develop sickle cell disease and not survive to adulthood. A person who carries the sickle cell trait (AS) will have a survival advantage against malaria in comparison to someone who doesn’t (AA).

We consider a simple model in which deaths occur due to either malaria ($M$) or to other causes, denoted $P$. The deaths that we are concerned with are those between birth and adulthood, which is taken to be the age at which children are produced. The number of adults from a cohort of newborns will be given by

$$\text{Surviving adults} = (1-M)(1-P) \text{ newborns}$$

Throughout the analysis, we will assume that the probability of dying from non-malaria causes, $P$, is the same for individuals with the AS and AA phenotypes.\footnote{Strictly speaking, $P$ is the probability of dying of some non malaria cause \textit{conditional} on not dying of malaria.} The probability of dying from non-malaria causes is equal to one for the SS phenotype. The probabilities of dying from malaria, differ between AA and AS phenotypes. We designate these probabilities $M^{AA}$ and $M^{AS}$.
It is also useful to designate the relative survival rates of these two phenotypes, which we call $\beta$:

$$
\beta = \frac{(1-M^{AA})(1-P)}{(1-M^{AS})(1-P)} = \frac{1-M^{AA}}{1-M^{AS}}
$$

$\beta$ is the probability of a non-carrier living into adulthood relative to the probability of a carrier living into adulthood. The smaller is $\beta$, the larger is the advantage of the AS phenotype. A value of $\beta = 1$ would indicate that there is no advantage to carrying the sickle cell gene.

The values for $M^{AA}$ and $M^{AS}$, and thus for $\beta$, will depend on both the disease environment and the state of medical technology. For example, in a place where there are no malaria mosquitoes, $M^{AA}$ and $M^{AS}$ will both be equal to zero, and $\beta$ will be equal to one. Clearly, the availability of modern medical care should mean that both $M^{AA}$ and $M^{AS}$ are lower today than they were in the past. However, it is not clear a priori which mortality rate would be reduced by more.

**Relative Survival in Modern Populations**

Although our interest is in asking what role malaria played historically, it is of interest to see what information is available about relative survival today. Nelson and Williams (2006) report that the prevalence of the the sickle cell trait in West Africa rises from 20-24% in newborns to 26-29% in adults. This is evidence for differential survival. We can convert this into a measure of relative survival, $\beta$, as follows: Taking the midpoint of each range (22% newborns and 27.5% of adults) and ignoring people who are SS, the data from Nelson and Williams imply that the ratio of AS to AA in newborns is 22/78 and in adults is 27.5/72.5. The equation relating the numbers of newborns and adults who are AA and AS is

$$
\frac{AA \text{ Adults}}{AS \text{ Adults}} = \frac{(1-M^{AA})(1-P) \times AA \text{ newborns}}{(1-M^{AS})(1-P) \times AS \text{ newborns}} \quad (3)
$$

Rearranging this equation:

$$
\beta = \frac{1-M^{AA}}{1-M^{AS}} = \frac{AA \text{ adults}}{AS \text{ adults}} \frac{AS \text{ newborns}}{AA \text{ newborns}} \quad (4)
$$

The implied value for from the Nelson and Williams data is $\beta = .74$.

Another study that can be used to examine relative survival, reported in Motulsky (1964, table 2) examined the relative survival of over 15,000 children in the Congo in the 1950s, a time and place where modern treatments for malaria would have been relatively scarce. The study compared families where one parent was AS and one AA, on the one hand, to families in which both parents were AA, on the other. Since half of the children in the former group were carriers,
compared to none in the latter, one can back out the relative survival of AS vs. AA children. The
study found mortality from all causes of 24.0% among AS children vs. 27.4% among AA. The
implied value of $\beta$ is .955. Part of the explanation for the different estimates $\beta$ in the Congo vs.
West Africa may be a difference in the severity of malaria. In the Congo, the malaria ecology
index is 12.1; in West Africa it is generally in the neighborhood of 20.

3.2 Measuring Relative Survival in Historical Populations

We now turn to our main line of inquiry, which is using observed frequency of the sickle cell
trait to back out the severity of malaria. Let $\pi_t$ be the fraction of adults in generation $t$ who are
carriers (AS). We assume that no one born with SS lives into adulthood. Thus the fraction of
the adult population who are not carriers is $(1 - \pi_t)$. The fraction of alleles in the adult
generation that are S is simply $\pi_t$. Assuming that mating between carriers and non-carriers is
random, the fractions of children born who are (AA), (AS), and (SS) are $(1 - \pi_t)^2, \pi_t(1 - \pi_t)$,
and $(\pi_t)^2$, respectively. The difference equation for $\pi$, which relates prevalence among adults in
successive generations, is

$$\pi_{t+1} = \frac{\pi_t (1 - \pi_t) / 2}{\pi_t (1 - \pi_t) / 2 + \beta (1 - \pi_t) / 2} = \frac{\pi_t}{\beta + \pi_t (1 - \beta / 2)} \quad (5)$$

We solve for the steady state by setting $\frac{\pi_{t+1}}{\pi_t} = 1$:

$$\pi_{ss} = \frac{1 - \beta}{1 - (\beta / 2)} \quad (6)$$

This has the properties we would expect: the smaller is the $\beta$, that is the greater is the survival
advantage of being a carrier, the larger is the equilibrium fraction of the adult population that
will be carriers.

We can turn this equation around to infer the burden of malaria on survival based on the
prevalence of the sickle cell trait among adults:

$$\beta = \frac{1 - \pi_{ss}}{1 - \pi_{ss} / 2} \quad (7)$$
This equation says that if 20% of the adult population are carriers in a steady state, then $\beta = 0.89$, in other words that people with the AA were only 89% as likely to live to adulthood as those with AS.

**Current vs. Historical Prevalence**

The analysis above considers a population that is in equilibrium in terms of the selective impact of malaria and the prevalence of the sickle cell trait. Such a steady state presumably existed in Africa in the period before European contact. However, the only data on sickle cell prevalence available comes from observations over the last 60 years. One might worry that modern prevalence rates are not the same as those that held historically, because the health environment has been changing over time. To address this question, it is straightforward to use equation (5) to look at how prevalence $\pi$ changes in response to a change in relative survival ($\beta$).

As an example, we consider the case where there is initially a steady state of $\pi = 0.20$, and correspondingly $\beta = 0.89$. In generation 1, the value of $\beta$ is set to one, corresponding to the complete eradication of malaria, or the removal of the population to a place where the disease is not present. Table 6 shows the fraction of the population that will be carriers of the sickle cell trait after a set number of generations.

The table shows that the initial decline is very rapid, but that there is long tailing off once the prevalence gets sufficiently low. In principle, one could use this setup to “backcast” prevalence of the sickle cell trait by looking at populations that were removed from malarial areas at known times. Of course a precise backcast would rely crucially on knowing exactly how many generations removed a population was removed from a malarious region, which would be practically impossible. However such an exercise can serve as a good check on the general reasonableness of estimates. For example, 8% of African Americans are carriers of the sickle cell trait. Suppose we take as a rough estimate that the average ancestor of today's African ancestor of today's African Americans left Africa around 1750 -- a span of roughly 10 generations. One also has to deal with the issue of admixture with other populations. Putterman and Weil (2010, Appendix B) summarize literature on the fraction of African American heritage that is due to non-Africans, reporting 20% as a rough consensus figure. That admixture took place at unknown points in time over the last three centuries, but assuming that it all took place at the end, the prevalence of the sickle cell trait in non-admixed blood would be 10%, and this would be consistent with the prevalence in the source population being 20% or a bit more. Given the origins of most slaves in the United States on Africa's west coast, this is a reasonable match with the prevalence data. Putterman and Weil give the largest source countries as Angola (15.4%), Ghana (12.0%), Senegal (10.7%), Nigeria (7.1%), Gabon (7.1%), Sierra Leone (6.8%), Guinea-Bissau (6.4%), Cameroon (5.4%), Congo (4.7%), and Gambia (4.3%).

---

14 Testing predates modern technology for genetic analysis. Carriers of the sickle cell gene can be reliably diagnosed by taking a drop of blood and mixing in an agent that induces hypoxia, then examining the cells under a microscope.
In practice, it is not clear that this analysis of the dynamics of prevalence matters much, since there is little reason to believe that contact with Europeans did anything to reduce the impact of malaria in Africa until the second half of the twentieth century, which is when most of the measures of sickle cell

3.3 Measuring the Overall Burden of Malaria

The above analysis does not tell us the overall effect of malaria, for three reasons. First, the parameter $\beta$ only tells us the relative survival of carriers and non-carriers. For example, a value of $\beta = 0.80$ could be consistent with $M^{AA} = 20\%$ and $M^{AS} = 0$, but it could also be consistent with $M^{AA} = 60\%$ and $M^{AS} = 50\%$. The second reason that $\beta$ alone does not tell us the overall burden of malaria is that we must take into account the cost of sickle-cell disease itself. Finally, to learn about the total burden of malaria, it is necessary to have information on probabilities of death from causes other than malaria or sickle cell disease. Recall that the malaria mortality measures, $M^{AS}$ and $M^{AA}$, are the probabilities of dying of malaria if one does not die of something else. The higher is the probability of dying of something else, the lower, by necessity, is the probability of dying of malaria.

Of these three issues, the second (the direct burden to sickle cell disease) is easily addressed by looking at the fraction of adults who are carriers, which is the same data used to estimate $\beta$. Knowing this, it is easy to figure out the fraction of children who will suffer from sickle cell disease. Assessing the other two issues requires bringing to bear additional data.

As noted above, $\beta$ tells us the relative survival of AS and AA children, but not the absolute rates of survival. To get an estimate of the burden of malaria, one needs an additional piece of information on $M^{AS}$, $M^{AA}$, or their ratio. One can look at modern populations for some information, with the caveat that modern data on survival is not necessarily informative about survival in Africa prior to European contact, where both the disease environment and the level of medical care differed from today. Allison (2002, Table 2) reports results from an examination of 104 child malaria deaths from different countries in Africa, in which the weighted average prevalence of the sickle cell trait was 21%. Only one child examined had the sickle cell trait, which would suggest that the trait is almost completely protective against malaria death. However, a different set of investigations (Allison 2002, Table 1) that looked at severe $P. falciparum$ infections rather than deaths found a relative incidence of infections in AS that was 46% as high that for AA. Both sets of studies just described were conducted in the 1950s or very early 1960s. A larger and more recent study (Hill et al., 1991) examined children in The Gambia. Children who were severely ill with malaria were compared to a control group. The severely ill children had cerebral malaria or severe malarial anemia. Without treatment, most of the children in this group would have died. Among the severe malaria group, the frequency of the AS phenotype was 1.2%, while among the control group it was 12.9%. This implies that the relative risk of developing severe symptoms (and presumably dying without medical care) in AS
as compared to AA is 0.08. These two studies suggest that reasonable bounds on $M^{AS}$ are zero on the low end, and to assume that $\frac{M^{AS}}{M^{AA}} = 0.08$ on the upper end.

With estimates of $M^{AA}$ and $M^{AS}$, we are in a position to look at the overall costs of malaria. There are three components to this cost: deaths from malaria among children who are carriers of the sickle cell trait, death from malaria among children who are not carriers, and deaths from sickle cell disease. Table 7 shows the fractions of births that fall into each category, the death rate for each group, and the total fraction of child deaths (from malaria or sickle cell disease) that are due to each category. The overall fraction of children who die due to malaria and sickle cell disease is simply the sum of the three terms in the right hand column.

Table 8 does a more extensive analysis, considering different values of $\pi$, the prevalence of the sickle cell trait among the adult population. We consider values ranging from zero to 40%, which is the highest level observed among specific populations. For each row, the second column calculates the implied value of $\beta$, assuming that the prevalence represents a steady state (equation 7). The third column shows the fraction of newborn children who will die of sickle cell disease. The fourth and fifth columns show the value of $M^{AA}$, the malaria death rate for non-carriers, under the two different assumptions about the death rate for carriers discussed above, specifically that $M^{AS} = 0$ and that $\frac{M^{AS}}{M^{AA}} = 0.08$. The sixth and seventh columns show the overall burden of mortality from both malaria and sickle cell disease, once again for the two assumption about the death rate for carriers discussed above.

The table shows that, at least within the range of the two estimates we have available, the assumption regarding the degree of protection afforded to carriers of the sickle cell mutation is not very important.

The fraction of the overall burden that takes the form of sickle cell disease rises with prevalence. For example, when $\pi = 0.20$, roughly one-tenth of the overall burden is in the form of deaths from sickle cell disease, with the other nine tenths due to malaria cases. When $\pi = 0.40$, sickle cell deaths account for roughly 20% of the burden. It is also of interest to calculate the net benefit of the sickle cell mutation, that is, the level of mortality in the presence of the mutation relative to the case where it is absent. The level of mortality absent the mutation can simply be read from the fourth or fifth column of the table since in this case everyone would have the mortality rate of non-carriers. For example, in the case of 20% prevalence of the sickle cell trait, overall mortality due to malaria and sickle cell disease is 10%, but overall mortality due to malaria would be 11.1% if there were no sickle cell mutation. In areas of high malaria pressure,

---

15 Another study (Greenwood, Marsh, and Snow, 1991) examined children in Kenya, finding that the sickle cell trait was present in only 1.8% of children with severe malaria anemia but 3.9% of children with uncomplicated malaria. This finding confirms that the trait is more protective against severe malaria than against mild cases. However, because no data are given on the prevalence of the trait in the overall population, one cannot back out the relative risk of AA vs. AS.

16 Given a ratio of malaria mortality in the two groups, along with the equation for $\beta$, we can solve the for the group rates of malaria mortality. These are $M^{AA} = \frac{1-\beta}{1-\beta X}$ and $M^{AS} = \frac{X(1-\beta)}{1-\beta X}$, where $X$ is the ratio of $M^{AS}$ to $M^{AA}$.
the value of the sickle cell mutation was higher. In the worst afflicted areas, the sickle cell trait reduced malaria related mortality by 20%.

Mortality from Non-Malaria Causes

As mentioned previously, the above calculation is incomplete in that it only describes the probability of dying of malaria conditional on not dying of something else. To know the absolute burden of malaria, one must know the death rate from other conditions, or alternatively the total death rate.

Systematic data on life expectancy in Africa is widely available only starting in the 1950s. For the period 1950-55, the United Nations estimate of life expectancy at birth in sub-Saharan Africa is 37.8 years (United Nations, 2009). Acemoglu and Johnson (2007) date the beginning of the “international epidemiological transition,” driven by more effective public health measures, the discovery of new chemicals and drugs, and international interventions to 1940. Although the transition came late to Africa, it is very likely that the 1950-55 figure represents an improvement relative to previous decades. Clearly after 1955 the pace of change was rapid. The UN estimates that life expectancy in sub-Saharan Africa rose by two years in each of the subsequent five year periods. Further evidence that health improvements were already underway by 1950 comes from data on total population size. Africa's population grew at a rate of 1.0% per year between 1900 and 1950, compared to a growth rate of 0.2% per year over the previous century (United Nations, 1999).

Data for the period prior to 1950 are very sparse. Acemoglu and Johnson (2007, Appendix C) pull together disparate sources to present a few estimates for the period before 1950. These are, Angola in 1940: 35 (both sexes); Mozambique in 1940: 45 (both sexes); Ghana in 1948: 38 (both sexes); and Mauritius, 1942-46: 32.3 (male) and 33.8 (female). Riley (2005) estimates that prior to the "health transition" (he uses a different definition than Acemoglu and Johnson) that began in Africa in the 1920s, life expectancy at birth averaged 26.4 years (this is the mean of 12 estimates, which range from 22.5 to 31.0.) In Asia, life expectancy prior to the health transition, which started there between 1870 and 1890, was 27.4. In Europe the transition started in the 1770s, and prior to it life expectancy was 34.3.

Riley comments that available estimates of African mortality prior to the health transition all come from European colonies in Africa. There is a reasonable basis for thinking that life expectancy may have been higher prior to colonization, the arrival of Arabic speaking merchants, and the dislocations produced by the slave trade. Unfortunately, little information for this period are available.

Steyn (2003) examines mortality in the pre-colonial period in northern South Africa thorough an examination of skeletal remains. She estimates life expectancy in the period 1000-1300 AD at

---

17 Some representative values are Angola: 27 years in 1940; Egypt: 30-33 years in the 1930s; Ghana: 28 years in 1921;Kenya: 23.9 years in the 1930s; South African black population: 38.1-40 years in 1935-40; Tunisia: 28.8 in the 1920s; Uganda: 23.9 in the 1930s; Zimbabwe: 26.4 in mid 1930s.
23.2, with the probability of surviving to age 20 being 48%. Remains for the post-1830 period show a slight decline in life expectancy after the expansion of European influence.

In the UN Model Life Tables for developing countries (1982), life expectancy at birth of 35 (general model, for females) implies that 61.4% of girls will survive to age 20 (and only 53.0% to age 30). The majority of premature deaths are concentrated at very low ages: 22.5% of girls would not reach age 5. Unfortunately, the model life tables that are available, to the extent that they reflect African data at all, certainly do not reflect the pattern of age-dependent mortality that existed in the period before the modern health environment of both disease and treatment was in place. It is likely that the pattern of mortality, and in particular the ratio of deaths at different ages, differed from what is observed today, but a priori there is no basis for guess the nature of this difference. Allowing for the fact that life expectancy at birth was almost certainly lower than 35 (the lowest value available in the UN model tables), and that "childbearing age" is not a single number but a range, a reasonable estimate for a value to use in our model is that 50% of births did not reach childbearing age.

To use the above number to calculate a burden of malaria, we have to pick a specific level of malaria prevalence. In the model laid out above, the level of malaria mortality \( M \) is unrelated to the level of non-malaria mortality \( P \). If this is correct, then areas with higher malaria burdens would have had lower overall survival rates. No data exist to test this assumption. For illustration, we will consider a case where prevalence of the sickle cell mutation is 20% and assume that total mortality before adulthood is 50%. From Table 8, the value of \( M \) is 10% (taking the second to last column, for mathematical simplicity, and also including both malaria and sickle cells deaths as part of \( M \)). Thus

\[
.5 = (1-M)(1-P) \tag{8}
\]

and \( P=0.444 \). To get a count of the fraction of children dying of malaria we now have to deal with an ambiguity in the model of survival with which we started. Specifically, the model implies that a fraction \( (1-M)(1-P) \) of children will survive both malaria and other conditions, but it is less clear about what those who don't survive die of.\(^{18}\) If malaria mortality comes before that from other conditions, then a fraction \( M \) will die of malaria and a fraction \( (1-M)P \) will die of other causes; if other conditions come first, then a fraction \( P \) will die of other causes and \( (1-P)M \) will die of malaria. The truth is obviously somewhere in the middle, although malaria mortality is probably more weighted toward the earliest part of life than other causes. For lack of any firm data, we simply assume that deaths due to malaria and "other" had equal time profiles. Specifically, of the fraction \( MP \) of children who don't survive either malaria or "other," we assume that \( M/(M+P) \) die of malaria and \( P/(M+P) \) die of something else. The total fraction of children who die of malaria is thus

\[^{18}\) The fraction who die is \( M+P-MP \). This can be rewritten as \( M(1-P)+P(1-M)+MP \), where the first term is children who died of malaria but would not have died otherwise, the second term is children who died of something else but would not have died of malaria, and the third term is children who died of one but would have died of the other.\]
For the values of $M$ and $P$ just given, this implies that 6.4% of children would have died from malaria, and malaria would have contributed 12.8% to pre-adult mortality.

Table 9 conducts this same calculation for the range of values of sickle cell prevalence, $\pi$, considered above. Rather than assume a constant overall pre-adult mortality rate (which would have the unlikely implication that non-malaria mortality fell one-for-one with rises in malaria mortality), we hold the value of $P$ constant at the value derived above (0.444) which is what would hold if the pre-adult death rate were 20% when sickle cell prevalence was equal to 20%. The table shows that, going from an area where malaria was absent ($\pi = 0$) to one of reasonably high prevalence ($\pi = .20$), the overall pre-adult death rate rises from 44% to 50%. In the worst afflicted areas ($\pi = .40$), the probability of pre-adult death was 56% and fully 25% of deaths were attributable to malaria and sickle cell disease.

### 3.4 Comparison to Modern Malaria Mortality Rates

The death rate from malaria derived from the historical data can be compared to modern rates. As mentioned above, for the WHO AFRO region, the under-five death rate from malaria is 0.59% per year. Multiplying this number by five gives an approximation to the probability of dying from malaria in the first five years of life, which is very close to the probability of dying from malaria before reproductive age. Although we cannot (at this point) assign values of sickle cell prevalence ($\pi$) to particular countries, it appears that the malaria death rate in the pre-European contact period was about twice as high.

A different and somewhat cleaner way to do the comparison of historical malaria mortality to mortality today is to focus on the variable $M$ itself, that is, the probability of dying of malaria conditional on not dying of something else. This is the measure constructed from historical data on prevalence, and although it is not generally reported, it can be constructed from modern data as well. Consider the case of Nigeria, which is a very heavily afflicted country. Annual malaria deaths for children under five are estimated to be 8.8 per thousand, or 0.88%. This implies that roughly 4.4% of children will die of malaria before their fifth birthday. The life table for Nigeria for 2006 shows that the probability of woman surviving to age 25 is approximately 0.75. Thus making the same assumptions about the timing of deaths from malaria and other causes as we made above, there are two equations:

\[
.75 = (1-M)(1-P) \tag{10}
\]

\[
0.044 = M(1 - P) + \frac{PM^2}{M + P} \tag{11}
\]

Solving these yields $M = 0.054$. Measured this way -- as the probability of dying of malaria conditional on not dying of something else -- the burden of malaria is far lower in Nigeria today.
than it was in countries with a sickle cell prevalence of 20% in the historical period. The implied value of \( P \) is 0.207. If we assume that Nigeria has a sickle cell prevalence of 20%, the implied values of \( M \) and \( P \) for the historical period were .10 and .444, respectively. Thus the probability of death from other causes has fallen slightly more than the probability of death from malaria, although the magnitudes of the two declines are close. Certainly, given the uncertainty in so many aspects of measurement, one cannot reject the possibility that the percentage decline in malaria mortality has been the same as the decline in other types of mortality.

4. Assessing the Importance of Malaria to Early African Development

Now that we have an estimate of how large a burden malaria imposed in terms of survival, we can turn to the question of whether malaria had a significant impact on African development. We divide our discussion into two parts, looking first at the effects of malaria deaths (mortality) and then at the effects of ill health among those who survive the disease (morbidity).

4.1 Direct Effect of Malaria Mortality

Current analyses of the burden of disease focus on measures such as years of life lost or disability free life years lost. From this perspective, the death of a young child is particularly costly because he or she had so many potential life years. The ethical considerations regarding the allocation of scarce lifesaving resources, and implicitly the cost of death and disease experienced at different ages, are quite complex (see Persad, Wertheimer, and Emanuuel, 2009). In assessing the role that disease played in affecting development historically, however, it seems reasonable to take a purely instrumental view of life and health, in which the primary considerations are how much society has invested in an individual and that individual's potential to produce services for society in the future. Under this view, the most costly death is that of a young adult, who has consumed a good deal of resources (food, childcare, education), and who has many years of potentially productive labor. The death of an infant or small child, by contrast, is far less costly, because fewer resources have been invested; and the death of an old person is similarly less costly because less productive potential is lost. Under this interpretation, malaria deaths are relatively low cost, with the exception of deaths of women in their first pregnancies, who are near the peak of their value as assets to society in terms of the balance between resources invested in them and services they can deliver.

To formalize this idea, consider a simple model of production and consumption with individuals of different ages. Let \( c_i \) be the consumption of an individual of age \( i \), and similarly \( w_i \) be labor income. For now, we ignore income from non labor sources, and assume that there is no storage of output between period. Let \( N_i \) by the number of people in age group \( i \). The social budget constraint is

\[
\sum_{0}^{T} N_i (w_i - c_i) = 0
\]
where $T$ is the maximum lifespan. We assume that consumption at each age is determined by two things: a consumption level of individuals at some benchmark age (for example, prime age adults), which we call $\bar{c}$, and some age-varying relative consumption coefficient $\tilde{c}_l$.

$$c_l = \tilde{c}_l \bar{c}$$

The values of $\tilde{c}_l$ presumably reflect both changing biological needs for consumption over the course of the life cycle as well as the arrangements by which consumption is divided up among different groups in society. One would not necessarily expect the pattern of consumption to be the same in all societies at all times. However, as discussed below, available data do not vary all that much.

We assume that wages at each age group are determined in a similar fashion:

$$w_l = \tilde{w}_l \bar{w}$$

where $\bar{w}$ is the return to some standard unit of labor. Again, the pattern of relative labor input across age groups reflects both biological differences and differences among societies in the economic value of different characteristics (for example, strength vs. wisdom vs. manual dexterity).

Combining the above equations:

$$\bar{c} = \bar{w} \frac{\sum_{l=0}^{T} N_l \tilde{w}_l}{\sum_{l=0}^{T} N_l \tilde{c}_l}$$

To use this equation to assess the effect of disease mortality on the consumption benchmark, we need estimates of the consumption and income profiles as well as an estimate of the effect of disease on the age structure of the population.

**Consumption and Income profiles**

A number of sources provide data on the life cycle profiles of consumption and labor input. Mueller (1976) synthesizes data from nine societies practicing what she calls “peasant agriculture,” by which she means agricultural systems which use primarily traditional methods of cultivation, small landholdings, and low capital inputs. South and Southeast Asia are primarily what she has in mind. The profiles are shown in Figure 11. Note that Mueller’s data on labor input apply to production of output as it would appear in measured GDP but exclude home production. Much of the latter is done by women, so in her data, productivity by prime age women is only 30% of the level of similarly aged males. Because we are interested in both sorts
of output, we focus on the male profiles. For consumption, Mueller provides two profiles in which consumption of people at each age (and of each gender) is compared to males aged 20-54. For the medium consumption profile, children age 0-4 have a value of 0.32; for the low consumption profile, the value is 0.12 (prime aged women get a value of 0.80). We use the medium profile.19

Population Age Structure

The other piece of data used in the equation above is \( N_i \), the number of people in each age group. In general, this will be a function of both the probability of survival to each age and the history of births or population growth. However, for the long historical periods that we are considering, population growth must have been very close to zero, which implies a constant number of births per year. We can thus approximate the age structure of the population \( N_i \) with the fraction of survivors at each age from the life table.

Our approach to assessing the role of disease in affecting consumption possibilities is to start with a baseline life table and then consider alternations that would result from eliminating or adding particular sources of mortality. To construct these alternative life tables, we use information on the age of death from different diseases. As our baseline, we use the United Nations (1982) model life table for a population with life expectancy at birth of 35 years (male and female combined).

This model can be used to analyze the cost, in terms of consumption, of deaths at different ages. Consider an increase in age specific mortality of an amount \( x \) at age \( s \), with no change in age specific mortality at other ages. We consider steady state population age distributions, so that the change in mortality at age \( s \) will affect the number of survivors at all higher ages. The level of consumption in the new steady state will be

\[
\bar{c}' = \frac{\sum_j N_j \bar{w}_j - x \sum_j N_j \bar{w}_j}{\sum_i N_i \bar{c}_i - x \sum_i N_i \bar{c}_i}
\]

19 We also looked at data from two other sources. First, Lee and Mason (2009) look at data from four contemporary hunter-gatherer societies originally studied by Kaplan (1994) and Howell (in press). The underlying data are in terms of calories collected and consumed. Their data are quite similar to the profiles in Mueller. Second, Lee and Mason also look at data for the four poorest countries that are part of the National Transfer Account project: Kenya, Philippines, Indonesia, and India. Income in this case is labor income, including unpaid family labor, and pertains to both men and women. These data differ from the other two sources primarily in showing a decline in income of the elderly that is not present in the peasant agriculture or hunter-gatherer data.
The proportional change in consumption is

\[ \frac{\bar{c}' - \bar{c}}{\bar{c}} = \left( 1 - \frac{x \sum_{i=0}^{T} N_i \bar{w}_i}{\sum_{i=0}^{T} N_i \bar{w}_i} \right) \left( 1 - \frac{x \sum_{i=0}^{T} N_i \bar{c}_i}{\sum_{i=0}^{T} N_i \bar{c}_i} \right) - 1 \]

Figure 12 shows the change in \( \bar{c} \) resulting from a one percent increase in mortality at different ages. By assumption, the effect of changes in mortality at birth is zero, since higher infant mortality does not change the age structure of the population.\(^{20}\) The loss to society from deaths rises with age, and peaks at the age in which income exceeds consumption. For example, a one percentage point rise in the mortality probability of 15 year olds leads to 0.2% decline in steady state consumption. The consumption impact of mortality declines as expected lifetime and productivity decline. For individuals at or nearing an age where consumption exceeds income, there is a social gain in terms of consumption from an increase in the death rate, although in this data the gain is quite small, both because consumption does not exceed income by much at any age and because the fraction of people who are expected to live to old age is small to begin with.

To apply this framework to different diseases, one needs a profile of the mortality effect of the disease at different ages. For malaria, we use age-group death rates from Murray and Lopez (1996).\(^ {21}\) To model the effect of malaria on the historical life table we proceed as follows. We start with the UN life table (life expectancy of 35) discussed above. We take this as a benchmark. To this we add additional mortality at each age proportional to the current age profile of malaria. Specifically, we take the current death rates from malaria and multiply them by a scaling factor, and add these death rates to the death rates in the UN life table. The scaling factor is chose to match the magnitude of the change in malaria deaths we want to model. Specifically, above for the case of going from zero malaria to the case where sickle cell prevalence is moderately high (\( \pi = .2 \)), we showed that the total pre-adult death rate rose from 44% to 50%. In the UN data, the probability of living to age 28 is .555, which is close enough to a 44% pre adult death rate (In the model "adult" is the age at which children are produced. In reality, there is no single age that is appropriate, but 28 is roughly in the middle of the relevant range.) The scaling factor required to set survival to age 28 equal to 50% is 2. Figure 13 shows the survival curves for the baseline case and the case with high malaria.\(^ {22}\) Life expectancy falls by 3.5 years from this additional malaria mortality.

\(^{20}\) One way to improve the model would be to include a consumption cost and decrement to labor input associated with pregnancy. This would involve setting "age zero" to be the year before birth, and altering the consumption and wage profiles accordingly. In this case, the loss to society from mortality at birth would no longer be zero, although it would still be relatively small.

\(^{21}\) These are 0.00559 for ages 0-4; 0.00042 for ages 5-14; 0.00033 for ages 15-44; and 0.00036 for ages 45-59.

\(^{22}\) Of course the UN table already includes deaths due to malaria. Another approach would be to take as the baseline a current life table stripped of malaria mortality, and then to add malaria mortality back in. This procedure would
This increase in malaria mortality lowers steady state consumption by only 0.63%. This is obviously an extremely small effect for such a large change in mortality. The reason is clear once one has seen Figure 12 and noted that malaria deaths are concentrated at very low ages. In this calculation, deaths beyond age five account for only 1/3 of the reduction in life expectancy due to malaria, but for 2/3 of the economic cost of the disease.

The procedure above can be applied to any disease or other cause of death. For a comparison to malaria, we look at smallpox, another one of the most important diseases in shaping human development. Smallpox is a virus spread from person to person by direct contact, aerosol droplets, or contact with infected objects. The case fatality rate for the disease in the absence of modern treatments is roughly 25-30%, with many survivors left scarred and some blinded. In areas of high population density the disease was a constant presence, so that cases were concentrated among the young, although even these locations (that is, big cities) were nonetheless subject to periodic epidemics. Areas of lower population densities, which could not support the disease in its endemic form, were more susceptible to periodic epidemics in which a large population of individuals without any immunity were stricken.

Smallpox had appeared in China, India, and Egypt well before the birth of Christ, and was definitely established in Europe by the end of the sixth century. Climate and low population density in Africa were not as conducive to the spread of the disease. Smallpox was not endemic to Africa south of the Sahara (with the exception of Ethiopia) prior to the beginning of the second millennium. The disease was spread by Portuguese traders along the west coast of the continent in the sixteenth and seventeenth centuries, and by Arab traders on the East coast in the 16th century (Hopkins, 2002).

Unlike malaria, there is no genetic fingerprint that one can look to in order to gauge the mortality pressure of smallpox. Information from various sources makes it clear that the disease was an important killer, however. In London and Geneva from 1580-1780, smallpox deaths averaged roughly 6% of total deaths (Hopkins, 2002, figure 5). According to the World Health Organization, in the 18th century smallpox killed every tenth child born in Sweden and France and every seventh child born in Russia. Behbehani (1983) reports estimates that at the end of the 18th century, smallpox was killing 400,000 Europeans per year. Maddison (2006) gives the total population of Europe (Western and Eastern) plus European Russia as 241 million in 1820, slightly after the end of the 18th century. This implies a death rate of 0.17% of the population every year; with a life expectancy of 40 years, this would imply a roughly 6.6% chance of dying from smallpox on a lifetime basis.

yield very similar results.

23 This case fatality rate applies to variola major. The case fatality rate for variola minor is less than one percent. Infection with either form provides lifelong protection against both.

24 http://www.who.int/mediacentre/factsheets/smallpox/en/
25 However, it is not clear that the European figure also includes Russia. If it does not, then the death rate would be higher.
There is also useful information on the age of death of smallpox victims [to be added]

The above analysis shows that, at least in terms of the extra costs associated with raising children who were subsequently going to die of malaria, the effect of the disease on the standard of living in Africa should have been relatively low. However, it is possible that there are other channels through which high malaria mortality may have mattered. Death from disease may impose an obstacle to economic development beyond the instrumental cost discussed here. Stone (1977) argues that in the face of high child mortality, pre-modern Europeans avoided forming emotional bonds with their young children, even to the point of giving the same name to two living children, expecting only one to survive. In such an environment, one might also expect to see less investment of tangible resources in the human capital (both health and education) of their children. Thus even if high child mortality from malaria were not important in the strict sense of draining resources from the economy, it might have contributed to reduced investment in children. Several recent theories of economic growth have stressed the importance of such investments to long-term development.

*The Malthusian Perspective*

The analysis above takes the perspective that deaths are costly in social terms, since they represent the loss of potential productive workers. An alternative view points to a flip side of deaths, which is that they decrease the size of population and thus lower pressure on fixed resources.

Consider a simple Malthusian model, as represented in the two panels of Figure 14. In both panels, the horizontal axis measures the standard of living. In a modern economy, this would be GDP per capita, but the model should be applicable even in contexts where there are no markets or prices with which to measure GDP. In the top panel, the vertical axis measures the size of the population. The downward sloping line in the panel represents the effect of higher population size in reducing living standards through the crowding of fixed resources such as land. This crowding could take the form of shorter fallow periods, higher pressure on grazing land, or increased levels of animal disease, all of which would reduce per-capita yields.

The bottom panel shows the relationship between the growth rate of population and the standard of living. We assume that there is an upward sloping relationship between these. In the classical Malthusian model, this positive relationship comes from either the "preventive check" (reduced desired fertility) or the "positive check" (higher mortality or lower fecundity) when income is low. Another mechanism that will produce very similar dynamics is if higher population density raises the mortality rate by facilitating the transmission of disease. In the panel, this upward sloping line is drawn twice, at a higher level to represent a place where the health environment is good, and at a lower level to represent a place where the health environment is bad. For a given standard of living, population growth will be higher in a healthier environment. In either case, the equilibrium of the system is reached at the level of population which implies a standard of living that in turn implies zero population growth.
One of the defining characteristics of the Malthusian model is that it produces a steady state in which, as long as the quantity of land is constant the the production technology is fixed, the growth rate of population is zero and the population size is constant (or, more realistically, the population size and standard of living are mean reverting in response to stochastic shocks). The near constancy of human populations over long historical eras suggests that such a mechanism must have been operating in most places for most of human history (Galor and Weil, 2000). Although there are clearly times when Malthusian constraints were lifted (such as the peopling of the Americas in pre-historic times), a simple analysis of the effect of compounding shows that most of the time, average population growth rates must have been quite close to zero. Without a homeostatic model of the Malthusian sort, it is vanishingly unlikely that such near constancy of population would obtain.

As the figure shows, a worse health environment leads to lower population and a higher standard of living in equilibrium. Thus malaria would have actually raised the standard of living in Africa. While counterintuitive, this conclusion is in line with the arguments in several recent analyses of growth. Voightlander and Voth (2009) argue that in Europe before the industrial revolution, income rose as a result of increased mortality due to plague, urbanization, and warfare, and that this rise in income was instrumental in knocking the continent onto the path of industrialization. Acemoglu and Johnson, in their analysis of the worldwide epidemiological transition of the mid 20th century, find that countries that experienced greater increases in life expectancy saw slower growth in income per capita -- a result that they attribute to the Malthusian effect of population growth. And Young (2005) claims that higher mortality and lower fertility due to HIV in South Africa will more than compensate for declines in worker productivity due to the disease, so that income per capita will actually rise as a result of the epidemic. Ashraf, Lester, and Weil (2008) examine the Malthusian channel as part of a broader analysis of how health improvements affect growth. They find that higher survival indeed pushes income lower through the Malthusian channel (as well as through capital dilution), although this effect is more than compensated for in the long run by higher labor productivity due to better health.

Whether this sort of Malthusian model applies to Africa in the period before European contact is unclear. Several authors stress that land shortage was not generally a problem in Africa. For example, Herbst (2000) argues that the abundant land was a persistent characteristic of African economies, as a result of which it was the control of people, rather than territory, that was of primary interest to rulers.

However, historians are often too quick to dismiss the Malthusian model without having another explanation for the near constancy of population over time. It is conceivable that, rather than land being the binding resource, the role of equilibrating population size was played by density-dependent disease. This would be consistent with the often-expressed view that labor, rather than land, was the binding constraint faced by African economies. If so, the pressure from disease must have been extraordinarily high indeed, however, since human capacity to reproduce is very high. The levels of malaria indicated by the prevalence of the sickle cell trait would not
be sufficient to hold down population growth in the absence of a land constraint, although it is possible that malaria in combination with other diseases did so.

It is also possible that malaria did indeed raise the standard of living in the short run by reducing population density, but at the same time low population density (due both to malaria and low agricultural productivity) prevented the development of large agglomerations which in other places were the locus of specialization, gains from trade, and technological advance.

### 4.2 Economic Effects of Malaria Morbidity

To pursue the question of how much malaria affects labor input of adults, we use data from the World Health Organization on Years Lost to Disability (YLDs) from particular diseases. A country’s YLD for a given disease is constructed as:

\[
YLD = I \times DW \times L
\]

where \( I \) is the number of incident (newly-arising) cases in a period, \( DW \) is the disability weight attached to the disease, and \( L \) is the average duration of the disease until remission or death. The crucial parameter here is the disability weight, which is intended to be a cardinal measure of the severity of different diseases or impairments, on a scale from 0, indicating perfect health, to 1, indicating death. Disability weights are constructed by panels of healthcare providers and medical experts using a "person trade-off" protocol which establishes utility equivalences between years of life lived in different states of health. One year lived with a disability provides the same utility as \( (1-DW) \) years lived disability-free (Murray, 1996). Disability weights are therefore not primarily intended as a measure of labor supply. Nevertheless, these estimates provide at least some basis for comparing the effects of different diseases.\(^{26}\)

To give an example of the interpretation of YLD: in the data below, the average YLD from all causes for Africa is 0.12. This means that in the average year, the average man suffers disease episodes that deprive him of the equivalent of 12% of a year’s disability free life. This could mean being fully disabled for 12% of the year, 50% disabled for 24% of the year, and so on. It could also mean suffering an incident that leaves him, say, 1% disabled for the next 12 years.

Table 10 shows the YLDs by disease type and age (for males) for the WHO's AFRO region. YLDs are counted at the age in which a disease incident occurs. Thus, for example, the neurological sequelae of malaria are counted as years lost to disability in the age 0-4 age group, even though the actual years lost are spread through the individual's whole life. (Note that the "total" column is based on the distribution of the current population among age groups. Thus disabilities affecting the elderly, who are a very small percentage of the current African population, play a very minor role in determining the total figure.)

\(^{26}\) Some examples of disability weights are blindness (0.600), deafness (0.216), HIV (0.136), AIDS (0.505), tuberculosis sero-negative for HIV (0.264), severe iron-deficiency anemia (0.093), malaria episodes (0.172) and neurological sequelae of malaria (0.473).
The table shows that overall, malaria accounts for only 5% of total years lost to disability. Further, almost all of these lost years were due to incidents in the first five years of life. Some of the years lost to disability in this case affected adults, through the neurological sequelae of the disease, but much of the disability burden fell directly on children. Although disability (which in this case really just measures suffering) among children does not directly affect production, one could argue that it might have affected their accumulation of skills or human capital. In any case, however, the malaria burden is relatively small in comparison to that of other diseases.

As mentioned above, another potential channel through which malaria might have affected economic development is through shifting population away from potentially productive areas. Gallup and Sachs give examples of this occurring in Europe. In the context of Africa, we do not know of evidence that particular regions were not settled because of malaria, and there is certainly evidence of people settling in areas of tremendously high malaria transmission. Given that malaria deaths were concentrated among the very young, and that mortality in this group was very high from other causes in any case, it is hard to see how malaria would have had a great influence on settlement patterns.

Again, it is useful to compare the morbidity burden of malaria to other conditions. In the case of smallpox, the adult incidence of the disease was certainly lower than that of malaria, since after a single case (most often contracted as a child) individuals were immune for life. However, the physical damage done by smallpox could be severe. A commonly quoted estimate is that one-third of blindness in 18th century Europe was due to smallpox. If we had better data we could convert this into an actual estimate of the disability burden.

5. Conclusion

Sachs, Malaney, and Spielman (2004) write “how better to evince the power of the parasite than with a potentially lethal modification of the genetic code as a desperate Darwinian defense against the even more deadly ravages of malaria? Accordingly, it may be expected that a force strong enough to rewrite our DNA will rewrite many of the lives and economies that it touches.” In this paper we have tried to address this issue directly. That is, we have used the extent which malaria left its mark on the human genome to back out the severity of the disease’s impact, and then in turn we have tried to assess how large the economic impact of that disease would have been.

In areas of high malaria transmission, 20% of the population carry the sickle cell trait. Our estimate is that this implies that historically between 10 and 11 percent of children died from malaria or sickle cell disease before reaching adulthood. Such a death rate is roughly twice the current burden of malaria in such regions. Comparing the most affected to least affected areas, malaria may have been responsible for a ten percentage point difference in the probability of surviving to adulthood. In terms of its burden relative to other causes of mortality, malaria appears to have been perhaps slightly less important historically than it is today, although we
certainly can't rule out the decline in malaria mortality has been proportional to the decline in mortality from other diseases.

Thus, malaria imposed a heavy mortality burden. Did it hold back economic development? We find little reason to believe that it did. Examining the economic burden of malaria mortality in a simple life cycle model suggests that the disease was not very important, primarily because the vast majority of deaths that it caused were among the very young, in whom society had invested few resources. Our analysis of malaria morbidity, which is necessarily more speculative, also suggests a relatively minor effect of the disease on labor input, and thus on economic activity.

These model-based findings corroborate the findings of our statistical examination. Within Africa, areas with higher malaria burden, as evidenced by either malaria ecology or the prevalence of the sickle-cell trait, do not show lower levels of economic development or population density, either in the present or in the colonial era.
References


Ashraf, Quamrul, Ashley Lester, and David N. Weil, “When Does Improving Health Raise GDP?” *NBER Macroeconomics Annual 2008*.


Chenny, Shirley, Pascal St-Amour, and Desire Vencatchellum: “Slave prices from succession and bankruptcy sales in Mauritius, 1825–1827” in *Explorations in Economic History* 40 (2003) 419–442


Sachs, Malaney, and Spielman 2004, *Am J. trop hygiene and medicine*


United Nations, Department of Economic and Social Affairs, Population Division, (1982). *Model Life Tables for Developing Countries*


<table>
<thead>
<tr>
<th>Table 1: Malaria and Current Development</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable: Log real gdp per person 2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Malaria Ecology Index</td>
<td>-0.0577***</td>
<td>-0.0330***</td>
<td>-0.0354*</td>
<td>-0.0438***</td>
<td>-0.0377***</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Mean Malaria Ecology Index * Old World</td>
<td></td>
<td></td>
<td>0.0294</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old World Dummy</td>
<td></td>
<td></td>
<td>-0.526**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geographic Controls</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Continent Dummies</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Observations</td>
<td>160</td>
<td>160</td>
<td>48</td>
<td>160</td>
<td>132</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.603</td>
<td>0.631</td>
<td>0.329</td>
<td>0.628</td>
<td>0.648</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Geographic controls include distance to the nearest coastline, absolute latitude, percentage of the land surface area of the country that has any of the four Köppen-Geiger tropical climates, and mean elevation. To avoid perfect multicollinearity only a dummy for the Americas is included in column (4). Column (3) only includes African countries. Column (5) includes only countries from the Old World (defined as all countries minus the Americas and Oceania).
Table 2: Malaria and Population Density in Sub-Saharan African Countries

Dependent Variable: Log of Population Density in 1500 and 1950

<table>
<thead>
<tr>
<th></th>
<th>in 1500 (1)</th>
<th>in 1500 (2)</th>
<th>in 1500 (3)</th>
<th>in 1950 (4)</th>
<th>in 1950 (5)</th>
<th>in 1950 (6)</th>
<th>in 1950 (7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Malaria Ecology Index</td>
<td>0.0576***</td>
<td>0.0556</td>
<td>0.026</td>
<td>0.0102</td>
<td>0.049</td>
<td>0.0701</td>
<td>0.0948*</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.036)</td>
<td>(0.048)</td>
<td>(0.023)</td>
<td>(0.056)</td>
<td>(0.057)</td>
<td>(0.056)</td>
</tr>
<tr>
<td>Mean Land Quality Index</td>
<td>2.115</td>
<td>0.893</td>
<td>3.247**</td>
<td>2.721</td>
<td>2.604</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.484)</td>
<td>(1.554)</td>
<td>(1.541)</td>
<td>(1.904)</td>
<td>(2.063)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Malaria Ecology Index *</td>
<td>-0.0106</td>
<td>0.00764</td>
<td>-0.114</td>
<td>-0.235</td>
<td>-0.259</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Land Quality Index</td>
<td>(0.101)</td>
<td>(0.112)</td>
<td>(0.123)</td>
<td>(0.163)</td>
<td>(0.160)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Geographic Controls
Colonial Rule Dummies
Observations 46 45 45 48 47 47 47
R-squared 0.17 0.30 0.50 0.00 0.14 0.35 0.43

Robust standard errors in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Geographic controls include distance to the nearest coastline, absolute latitude, and percentage of the land surface area of the country that has any of the four Köppen-Geiger tropical climates. Colonial rule dummies are British, French, Spanish, and Portuguese.
Table 3: Malaria and Population Density


<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Malaria Ecology Index</td>
<td>0.0243* (0.013)</td>
<td>0.0244* (0.014)</td>
<td>0.0253* (0.013)</td>
<td>0.0250* (0.013)</td>
<td>0.0919*** (0.024)</td>
<td>0.103*** (0.025)</td>
<td>0.107*** (0.025)</td>
<td>0.107*** (0.025)</td>
<td>0.102*** (0.026)</td>
</tr>
<tr>
<td>Mean Land Quality Index</td>
<td>1.351** (0.567)</td>
<td>3.580*** (0.903)</td>
<td>3.048*** (0.852)</td>
<td>3.269*** (0.908)</td>
<td>3.278*** (0.916)</td>
<td>3.201*** (0.929)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Malaria Ecology Index * Mean Land Quality Index</td>
<td>-0.154*** (0.044)</td>
<td>-0.128*** (0.041)</td>
<td>-0.126*** (0.040)</td>
<td>-0.126*** (0.041)</td>
<td>-0.121*** (0.041)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Decade Dummies</th>
<th>Climate FE</th>
<th>Geographic Controls</th>
<th>Colonial Origin Dummies</th>
<th>Ocean Slave Trade</th>
<th>North Africa in Sample</th>
<th>Observations</th>
<th>R-squared</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Y Y Y Y Y Y Y Y</td>
<td>N N Y Y Y Y Y Y</td>
<td>N N N N N Y Y Y Y</td>
<td>N N N N N N N Y Y</td>
<td>N N N N N N N N</td>
<td>Y Y Y Y Y Y Y Y</td>
<td>206 206 206 206 206 206 206 206</td>
<td>0.018 0.08 0.346 0.365 0.403 0.488 0.504 0.504 0.459</td>
</tr>
</tbody>
</table>

Ordinary Least Square regressions. Robust standard errors in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively. The decade dummies cover the period 1840-1950 (omitted decade 1960). Climate FE is based on Köppen climate classification (omitted climate is subtropical highland -Cwb). Geographic controls include geodesic distance of the centroid of the historical homeland of each ethnic group from the nearest coastline, mean elevation of terrain, and absolute latitude. Ocean slave trade controls represent the number of persons of each ethnic group that were shipped during the Atlantic and Indian Ocean slave trades for the period 1400-1900.
Table 4: Malaria and Ethnic Prosperity

<table>
<thead>
<tr>
<th></th>
<th>(1) Mean Size of Local Communities</th>
<th>(2) Settlement Patterns</th>
<th>(3) Juridistional Hierarchy</th>
<th>(4) Subsistence Economy</th>
<th>(5) Intensity of Agriculture</th>
<th>(6) Prosperity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Malaria Ecology Index</td>
<td>0.113***</td>
<td>0.0282</td>
<td>0.0406**</td>
<td>0.0434***</td>
<td>0.00479</td>
<td>0.0637</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.018)</td>
<td>(0.016)</td>
<td>(0.014)</td>
<td>(0.017)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>Mean Land Quality Index</td>
<td>2.032*</td>
<td>1.188**</td>
<td>1.352**</td>
<td>1.105**</td>
<td>1.170*</td>
<td>3.271*</td>
</tr>
<tr>
<td></td>
<td>(1.208)</td>
<td>(0.562)</td>
<td>(0.578)</td>
<td>(0.530)</td>
<td>(0.611)</td>
<td>(1.905)</td>
</tr>
<tr>
<td>Mean Malaria Ecology Index</td>
<td>-0.0868</td>
<td>-0.0484*</td>
<td>-0.113***</td>
<td>-0.0642**</td>
<td>-0.034</td>
<td>-0.134*</td>
</tr>
<tr>
<td>* Mean Land Quality Index</td>
<td>(0.064)</td>
<td>(0.029)</td>
<td>(0.028)</td>
<td>(0.027)</td>
<td>(0.028)</td>
<td>(0.079)</td>
</tr>
</tbody>
</table>

Decade Dummies                  | Y Y Y Y Y                           | Y Y Y Y Y               | Y Y Y Y                     | Y Y Y Y Y               | Y Y Y Y Y                     | Y Y Y Y Y     |
Climate FE                      | Y Y Y Y                             | Y Y Y Y                 | Y Y Y                       | Y Y Y Y                 | Y Y Y                         | Y Y Y         |
Geographic Controls             | Y Y Y Y                             | Y Y Y Y                 | Y Y Y                       | Y Y Y Y                 | Y Y Y                         | Y Y Y         |
Colonial Rule Dummies           | Y Y Y Y                             | Y Y Y Y                 | Y Y Y                       | Y Y Y Y                 | Y Y Y                         | Y Y Y         |
Observations                    | 139                                | 380                     | 372                         | 409                     | 381                           | 136           |
Log pseudolikelihood            | -221.03                            | -503.20                 | -441.90                     | -506.38                 | -317.89                       | R2: 0.408     |

Ordered Probit Regressions. Last column presents OLS regression of first principal component of the 5 previous dependent variables. Robust standard errors in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively. The decade dummies cover the period 1840-1950 (omitted decade 1960). Climate FE is based on Köppen climate classification (omitted climate is subtropical highland -Cwb). Geographic controls include geodesic distance of the centroid of the historical homeland of each ethnic group from the nearest coastline, mean elevation of terrain, total land are and absolute latitude. Colonial rule dummies are British, French, Spanish, and Other.
Table 5: Sickle Cell Prevalence and Malaria Ecology

Dependent Variable: Prevalence of Sickle Cell Trait

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malaria Ecology Index</td>
<td>0.000996*</td>
<td>0.00517***</td>
<td>0.0178***</td>
<td>0.0178***</td>
</tr>
<tr>
<td></td>
<td>(0.0005)</td>
<td>(0.0009)</td>
<td>(0.0022)</td>
<td>(0.0028)</td>
</tr>
<tr>
<td>Malaria Ecology Index Square</td>
<td>-0.000477***</td>
<td>-0.000477***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.130***</td>
<td>0.0752***</td>
<td>0.0349***</td>
<td>0.0349***</td>
</tr>
<tr>
<td></td>
<td>(0.0090)</td>
<td>(0.0171)</td>
<td>(0.0111)</td>
<td>(0.0132)</td>
</tr>
<tr>
<td>Weighting</td>
<td>No</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Observations</td>
<td>670</td>
<td>670</td>
<td>670</td>
<td>366</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.005</td>
<td>0.19</td>
<td>0.303</td>
<td>0.341</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively. Weighting is based on the number of people tested in each location and each record. Column (4) uses a smaller sample with no duplications in location where the final figure for prevalence is the weighted average of the sickle cell trait prevalence using the number of people tested in each record as weights.
Table 6: Change Over Time in Fraction of the Adult Population who are Carriers

<table>
<thead>
<tr>
<th>Generation</th>
<th>Fraction Carriers</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.20</td>
</tr>
<tr>
<td>1</td>
<td>.181</td>
</tr>
<tr>
<td>2</td>
<td>.166</td>
</tr>
<tr>
<td>3</td>
<td>.153</td>
</tr>
<tr>
<td>5</td>
<td>.133</td>
</tr>
<tr>
<td>10</td>
<td>.100</td>
</tr>
<tr>
<td>15</td>
<td>.080</td>
</tr>
<tr>
<td>20</td>
<td>.066</td>
</tr>
<tr>
<td>25</td>
<td>.057</td>
</tr>
<tr>
<td>30</td>
<td>.050</td>
</tr>
</tbody>
</table>
Table 7: Components of the Cost of Malaria

<table>
<thead>
<tr>
<th>Group</th>
<th>Fraction of Births</th>
<th>Death Rate from Malaria or Sickle Cell Disease</th>
<th>Fraction of all Children who Die in Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Carriers (AA)</td>
<td>((1 - \frac{\pi}{2})^2)</td>
<td>(M^{AA})</td>
<td>((1 - \frac{\pi}{2})^2 M^{AA})</td>
</tr>
<tr>
<td>Carriers (AS)</td>
<td>(\pi \left(1 - \frac{\pi}{2}\right))</td>
<td>(M^{AS})</td>
<td>(\pi \left(1 - \frac{\pi}{2}\right) M^{AS})</td>
</tr>
<tr>
<td>Sickle Cell Disease (SS)</td>
<td>(\left(\frac{\pi}{2}\right)^2)</td>
<td>1</td>
<td>(\left(\frac{\pi}{2}\right)^2)</td>
</tr>
</tbody>
</table>

Table 8: Implications of Varying Sickle Cell Prevalence for Total Malaria Burden

<table>
<thead>
<tr>
<th>(\pi)</th>
<th>(\beta)</th>
<th>Fraction Dying of Sickle Cell Disease</th>
<th>(M^{AA}) assuming that (M^{AS} = 0)</th>
<th>(M^{AA}) assuming that (\frac{M^{AS}}{M^{AA}} = 0.08)</th>
<th>Overall Malaria and SC Burden assuming (M^{AS} = 0)</th>
<th>Overall Malaria and SC Burden assuming (\frac{M^{AA}}{M^{AS}} = 0.08)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.05</td>
<td>0.974</td>
<td>0.000625</td>
<td>0.026</td>
<td>0.028</td>
<td>0.025</td>
<td>0.0272</td>
</tr>
<tr>
<td>0.1</td>
<td>0.947</td>
<td>0.0025</td>
<td>0.053</td>
<td>0.057</td>
<td>0.05</td>
<td>0.0543</td>
</tr>
<tr>
<td>0.15</td>
<td>0.919</td>
<td>0.005625</td>
<td>0.081</td>
<td>0.088</td>
<td>0.075</td>
<td>0.0815</td>
</tr>
<tr>
<td>0.2</td>
<td>0.889</td>
<td>0.01</td>
<td>0.111</td>
<td>0.120</td>
<td>0.1</td>
<td>0.1086</td>
</tr>
<tr>
<td>0.25</td>
<td>0.857</td>
<td>0.015625</td>
<td>0.143</td>
<td>0.153</td>
<td>0.125</td>
<td>0.1357</td>
</tr>
<tr>
<td>0.3</td>
<td>0.824</td>
<td>0.0225</td>
<td>0.176</td>
<td>0.190</td>
<td>0.15</td>
<td>0.1628</td>
</tr>
<tr>
<td>0.35</td>
<td>0.788</td>
<td>0.030625</td>
<td>0.212</td>
<td>0.226</td>
<td>0.175</td>
<td>0.1899</td>
</tr>
<tr>
<td>0.4</td>
<td>0.75</td>
<td>0.04</td>
<td>0.25</td>
<td>0.266</td>
<td>0.2</td>
<td>0.2170</td>
</tr>
</tbody>
</table>
Table 9:

<table>
<thead>
<tr>
<th>π</th>
<th>β</th>
<th>M</th>
<th>Probability of Dying of Malaria or Sickle Cell</th>
<th>Total Death Rate (pre-adult)</th>
<th>Malaria and SC Deaths as Fraction of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.444</td>
<td>0.000</td>
</tr>
<tr>
<td>0.05</td>
<td>.974</td>
<td>0.025</td>
<td>0.014</td>
<td>0.458</td>
<td>0.032</td>
</tr>
<tr>
<td>0.10</td>
<td>.947</td>
<td>0.050</td>
<td>0.030</td>
<td>0.472</td>
<td>0.064</td>
</tr>
<tr>
<td>0.15</td>
<td>.919</td>
<td>0.075</td>
<td>0.047</td>
<td>0.486</td>
<td>0.096</td>
</tr>
<tr>
<td>0.20</td>
<td>.889</td>
<td>0.100</td>
<td>0.064</td>
<td>0.500</td>
<td>0.128</td>
</tr>
<tr>
<td>0.25</td>
<td>.857</td>
<td>0.125</td>
<td>0.082</td>
<td>0.514</td>
<td>0.159</td>
</tr>
<tr>
<td>0.30</td>
<td>.824</td>
<td>0.150</td>
<td>0.100</td>
<td>0.527</td>
<td>0.190</td>
</tr>
<tr>
<td>0.35</td>
<td>.788</td>
<td>0.175</td>
<td>0.119</td>
<td>0.541</td>
<td>0.220</td>
</tr>
<tr>
<td>0.40</td>
<td>.750</td>
<td>0.200</td>
<td>0.139</td>
<td>0.555</td>
<td>0.250</td>
</tr>
</tbody>
</table>

Table 10: Years Lost to Disability per capita, WHO AFRO Region

<table>
<thead>
<tr>
<th></th>
<th>0-4</th>
<th>5-14</th>
<th>15-29</th>
<th>30-44</th>
<th>45-59</th>
<th>60-69</th>
<th>70-79</th>
<th>80+</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Causes</td>
<td>0.174</td>
<td>0.068</td>
<td>0.133</td>
<td>0.134</td>
<td>0.141</td>
<td>0.126</td>
<td>0.115</td>
<td>0.103</td>
<td>0.122</td>
</tr>
<tr>
<td>Communicable,maternal,</td>
<td>0.112</td>
<td>0.028</td>
<td>0.040</td>
<td>0.037</td>
<td>0.021</td>
<td>0.015</td>
<td>0.013</td>
<td>0.010</td>
<td>0.046</td>
</tr>
<tr>
<td>perinatal, and nutritional conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>infectious and parasitic diseases</td>
<td>0.047</td>
<td>0.019</td>
<td>0.037</td>
<td>0.035</td>
<td>0.021</td>
<td>0.014</td>
<td>0.012</td>
<td>0.010</td>
<td>0.031</td>
</tr>
<tr>
<td>tuberculosis</td>
<td>0.000</td>
<td>0.001</td>
<td>0.002</td>
<td>0.004</td>
<td>0.002</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>HIV/AIDS</td>
<td>0.000</td>
<td>0.000</td>
<td>0.017</td>
<td>0.015</td>
<td>0.004</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.007</td>
</tr>
<tr>
<td>diarrhoeal diseases</td>
<td>0.004</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>childhood cluster diseases</td>
<td>0.007</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>malaria</td>
<td>0.027</td>
<td>0.001</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.006</td>
</tr>
<tr>
<td>tropical cluster diseases</td>
<td>0.001</td>
<td>0.009</td>
<td>0.011</td>
<td>0.010</td>
<td>0.009</td>
<td>0.004</td>
<td>0.003</td>
<td>0.003</td>
<td>0.008</td>
</tr>
<tr>
<td>Noncommunicable Diseases</td>
<td>0.047</td>
<td>0.019</td>
<td>0.066</td>
<td>0.074</td>
<td>0.108</td>
<td>0.105</td>
<td>0.098</td>
<td>0.089</td>
<td>0.056</td>
</tr>
<tr>
<td>Injuries</td>
<td>0.015</td>
<td>0.021</td>
<td>0.027</td>
<td>0.023</td>
<td>0.011</td>
<td>0.006</td>
<td>0.004</td>
<td>0.003</td>
<td>0.021</td>
</tr>
</tbody>
</table>

Source: Global Burden of Disease, 2002 revision.
Figure 1: Worldwide Distribution of Malaria Ecology

Malaria Ecology Index

- High: 38.1
- Low: 0
Figure 2: Malaria Ecology in Africa
Figure 3: Malaria Prevalence and Malaria Ecology

![Malaria Prevalence and Malaria Ecology graph]

coeff = 1158.7416, (robust) se = 163.12931, t = 7.1

Figure 4: Histogram Year of Estimation for Population Figure

![Histogram Year of Estimation for Population Figure]

Density

Year of Estimation for Population Figure in EA
Figure 5: Population Density by Ethnic Groups from EA (1967)
Figure 6: Population Density in West Africa from EA (1967)

Population Density (pop per sq km)
- 0.01 - 1.47
- 1.48 - 3.49
- 3.50 - 8.46
- 8.47 - 18.40
- 18.41 - 117.70

Figure 7: Histogram Year of Ethnographic Data
Figure 8: Distribution of Sickle Cell Prevalence Data and Malaria Ecology
Figure 9: Sickle Cell Gene Frequency from Piel et al (2010)


Figure 10: Sickle Cell Trait and Malaria Ecology
Figure 11: Consumption and Income Profiles
Figure 12: Change in Consumption Resulting from a 1% Increase in Deaths by Age
Figure 13: Change in Survival Probabilities due to Malaria
Figure 14: The Malthusian Model

- **Population Size**
  - Equilibrium population size with healthy environment
  - Equilibrium population size with unhealthy environment

- **Population Growth Rate**
  - Healthy Environment
  - Unhealthy Environment

- **Standard of Living**
  - Equilibrium standard of living with healthy environment
  - Equilibrium standard of living with unhealthy environment