



## Spatial Patterns of Infant Mortality in Mali: The Effect of Malaria Endemicity

A. Gemperli<sup>1</sup>, P. Vounatsou<sup>1</sup>, I. Kleinschmidt<sup>2</sup>, M. Bagayoko<sup>3</sup>, C. Lengeler<sup>1</sup>, and T. Smith<sup>1</sup>

<sup>1</sup> Swiss Tropical Institute, Basel, Switzerland.

<sup>2</sup> Medical Research Council, Durban, South Africa.

<sup>3</sup> Faculté de Médecine de Pharmacie et d'Otondo-Stomatologie, Université du Mali, Bamako, Mali.

Received for publication August 30, 2002; accepted for publication May 23, 2003.

A spatial analysis was carried out to identify factors related to geographic differences in infant mortality risk in Mali by linking data from two spatially structured databases: the Demographic and Health Surveys of 1995–1996 and the Mapping Malaria Risk in Africa database for Mali. Socioeconomic factors measured directly at the individual level and site-specific malaria prevalence predicted for the Demographic and Health Surveys' locations by a spatial model fitted to the Mapping Malaria Risk in Africa database were examined as possible risk factors. The analysis was carried out by fitting a Bayesian hierarchical geostatistical logistic model to infant mortality risk, by Markov chain Monte Carlo simulation. It confirmed that mother's education, birth order and interval, infant's sex, residence, and mother's age at infant's birth had a strong impact on infant mortality risk in Mali. The residual spatial pattern of infant mortality showed a clear relation to well-known foci of malaria transmission, especially the inland delta of the Niger River. No effect of estimated parasite prevalence could be demonstrated. Possible explanations include confounding by unmeasured covariates and sparsity of the source malaria data. Spatial statistical models of malaria prevalence are useful for indicating approximate levels of endemicity over wide areas and, hence, for guiding intervention strategies. However, at points very remote from those sampled, it is important to consider prediction error.

infant mortality; malaria; models, statistical; spatial analysis; variogram

Abbreviations: DHS, Demographic and Health Surveys; MARA, Mapping Malaria Risk in Africa; OR, odds ratio.

Malaria is an important cause of mortality among children in Africa, but the relation between malaria transmission intensity and child mortality remains controversial (1–5). A review of published studies of malaria-specific mortality shows some evidence that the highest mortality may be at intermediate transmission intensities (6). Rates of hospitalization with severe malaria in African children appeared to be highest at intermediate levels of transmission.

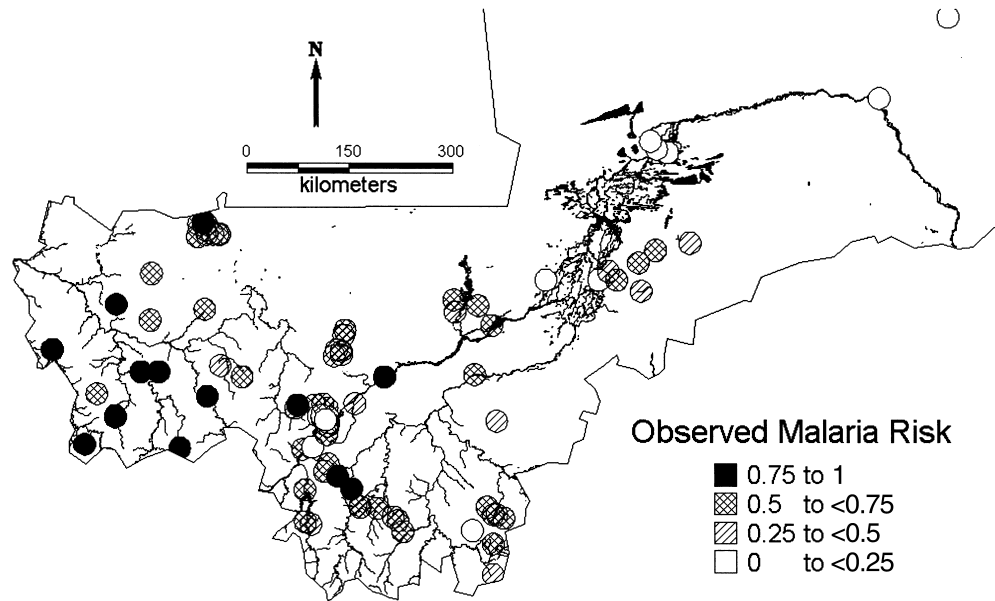
A difficulty with mortality studies is that the verbal autopsies used to assign a cause of death are not very reliable (3). Many deaths in malaria-endemic areas, assigned to other causes, are related to malaria infection (4). Moreover, low birth weight is an important risk factor for infant mortality and is known to arise because of both prematurity and intra-uterine growth retardation resulting from malaria infection of the mother during pregnancy (7). It follows that malaria may be a relevant risk factor for many deaths even when it is

not the immediate cause. Hence, it is as important to look at the relation of malaria endemicity to all-cause mortality as it is to look at its relation to malaria-specific deaths.

Smith et al. (5) linked published all-cause mortality rates and entomologic inoculation rates across Africa and found an increase in the infant mortality rate with these rates but no clear trend with the child (12–59 months of age) mortality rate. The major shortcomings of the study were the small number of sites compared and the fact that they were a convenience sample. Geographic variation in the factors independently affecting both malaria transmission and mortality (such as water availability) introduces ecologic confounding. Analyses linking site-specific mortality data with local malaria indices, which make appropriate adjustment for these ecologic confounders, are clearly needed.

There are a few data sets that allow these types of analysis. The databases we have used for this study are the Mapping

Reprint requests to Dr. P. Vounatsou, Swiss Tropical Institute, Socinstrasse 57, CH-4002 Basel, Switzerland (e-mail: penelope.vounatsou@unibas.ch).



**FIGURE 1.** Observed malaria prevalence in 34,800 children aged 1–10 years from the Mapping Malaria Risk in Africa surveys conducted in Mali between 1965 and 1998 at 101 locations. Rivers and lakes are also indicated.

Malaria Risk in Africa (MARA) and the Demographic and Health Surveys (DHS) databases. The MARA database consists of surveys recording malariologic information, with over 10,000 collected data points all over Sub-Saharan Africa, to date. It is currently the most comprehensive database on malariologic surveys in Africa (8). The DHS database, coordinated by Macro International, Inc. (Calverton, Maryland), provides nationally representative household surveys worldwide with large sample sizes of between 5,000 and 30,000 households, typically. It monitors indicators in the areas of demography, health, and nutrition.

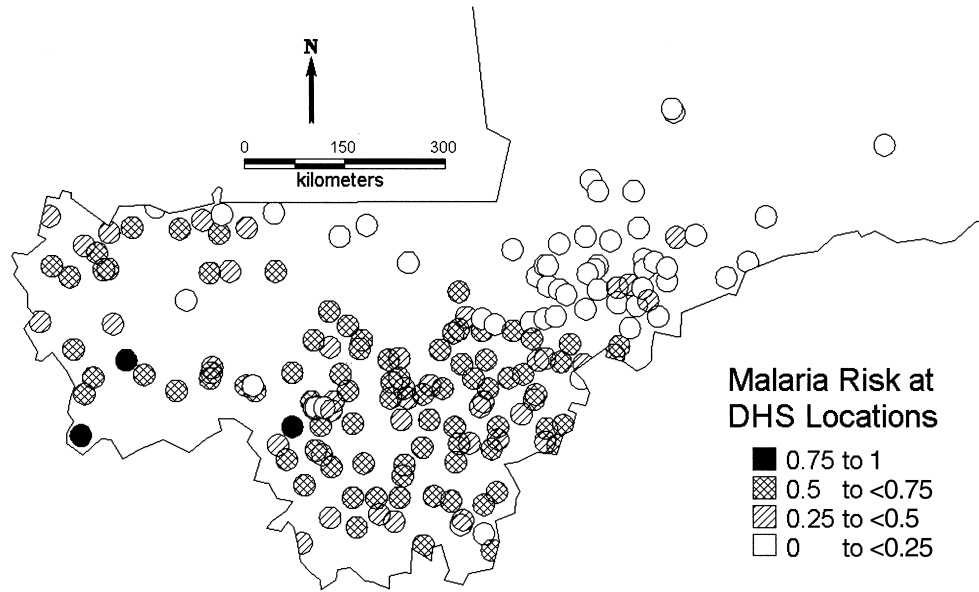
We linked the two databases using their site-specific data and developed a geostatistical model that enabled us to investigate spatial patterns of infant mortality risk, to assess its determinants, and to carry out an ecologic analysis examining the relation between infant mortality risk and malaria. We demonstrated the methodology by applying the technique to data collected from Mali. The savanna and Sahel zones of the country represent an appropriate setting for implementing such a methodology, because of the tendency for living conditions to become generally more difficult in the more northerly, drier areas, while malaria transmission is generally expected to be more intense in the wetter, southern areas. Therefore, we cannot properly assess the effect of malaria transmission on infant mortality without adjustment for potential confounders. We believe our approach of linking the two databases to carry out an ecologic analysis of infant mortality risk to be novel. Instead of a district-specific approach (9, 10), we modeled the data at the individual level, using a site-dependent correlation structure, to estimate the various effects without relying on data that are aggregated by administrative boundaries, which would have led to loss of information. This approach provides estimates of the effects

of various factors including malaria endemicity, taking into account confounding and spatial correlation. In addition, it allows us to produce maps of infant mortality risk adjusted for socioeconomic factors.

## MATERIALS AND METHODS

### Data sources

MARA (8) is an international collaboration initiated to provide a database and an atlas of malaria in Africa, by collating both published and unpublished results of malariologic surveys. Data on malaria endemicity were obtained from a model fitted to the MARA database by Kleinschmidt et al. (11). This was a spatial logistic model of malaria prevalence in children aged over 1 year and less than or equal to 10 years using the results of prevalence surveys conducted between 1965 and 1998 at 101 different locations in Mali. This age group was chosen because it has the highest prevalence and shows the clearest distinction between regional malaria endemicity patterns. In addition, most of the surveys were conducted for this age range and thus allowed inclusion of most of the data. The model included temperature, rainfall (12), the normalized difference vegetation index obtained from satellite data collected by the National Oceanic and Atmospheric Administration/National Aeronautics and Space Administration's Pathfinder Advanced Very High Resolution Radiometer Land Project, and distance from the nearest water body. The MARA survey sites are shown in figure 1. We used predictions of this model to obtain estimates of malaria prevalence at the DHS sample sites (figure 2). To avoid the linearity assumption between malaria prevalence and infant mortality, we converted these estimates to the following categories corresponding to different degrees



**FIGURE 2.** Estimated malaria prevalence at the 183 infant mortality sample locations in Mali obtained from the model of Kleinschmidt et al. (11). DHS, Demographic and Health Surveys.

of endemicity: 0–<0.25, 0.25–<0.50, 0.50–<0.75, and 0.75–1.00.

The DHS database is the most comprehensive database on child survival in Africa, compiled by a survey program in 23 countries funded by the US Agency for International Development and coordinated by Macro International, Inc. The aim of these surveys is to provide data for a range of monitoring and impact evaluation indicators within the population, health, and nutrition sectors. Birth histories since 1960 corresponding to 35,906 infants were extracted from the DHS carried out from November 1995 to April 1996 in Mali. Additionally, year of birth, residence, mother's education, infant's sex, birth order, preceding birth interval, and mother's age at infant's birth were extracted. Using location information provided by Macro International, we were able to geolocate the 183 distinct survey sites shown in figure 2 by using digital maps and databases, such as the "Africa Data Sampler" (13) and the "Geoname Digital Gazetteer" (14).

### Statistical analysis

Logistic regression models were fitted to infant mortality, using SAS version 8.2 software (SAS Institute, Inc., Cary, North Carolina) to identify significant socioeconomic, demographic, and birth-related covariates. The variables chosen were those analyzed by Coulibaly et al. (15) in their report that summarizes the results of the 1995–1996 DHS survey in Mali in relation to mortality. The mother's age at the child's birth and the date of birth of the child were included in all models. In addition, the following variables showing a significant bivariate association with infant mortality were selected for subsequent spatial multivariate

analysis: type of region, mother's education, sex, birth order, and preceding birth interval.

Bayesian hierarchical models were fitted to estimate the amount of spatial heterogeneity in infant mortality as well as associations between risk factors and infant mortality in the presence of spatial correlation. Ignoring this correlation would result in underestimation of the variance of the effects of risk factors (16). We used logistic models with village-specific random effects to assess geographic heterogeneity and the effects of different covariates. Some of the covariates used were at the individual level (socioeconomic, demographic, birth-related covariates) and some were at the village level (malaria prevalence category). Malaria prevalence was not observed at the locations of the mortality data but was estimated using the model by Kleinschmidt et al. (11). These estimates were categorized because explanatory analysis revealed a nonlinear relation between infant mortality and malaria prevalence. The cutoffs were chosen according to the frequency distribution of mortality data and on epidemiologic considerations. Spatial correlation was modeled by assuming that the random effects are distributed according to a multivariate normal distribution with a variance-covariance matrix related to the variogram of the spatial process (17). We used Markov chain Monte Carlo simulation (18, 19) to estimate the model parameters. Simulation-based Bayesian kriging was also applied to produce smoothed maps of mortality risk and of the variance of the map estimates (20). The deviance information criterion (21) was applied to assess the effect of malaria prevalence on infant mortality. The smaller the deviance information criterion values, the better the fit of the model. Further details of this modeling approach are given in the Appendix. The analysis was implemented using software written by the authors in Fortran 95 (Compaq Visual Fortran version 6.1) using

International Mathematical and Statistical Libraries (IMSL) numerical libraries (Visual Numerics, Inc., Houston, Texas).

## RESULTS

The median malaria prevalence estimated for the 183 locations was 49 percent, ranging from 4 percent to 82 percent; 16.9 percent of the 35,906 infants sampled died before completing their first year of life. This is untypically high compared with the recent census-based estimates of 12.3 percent (Population Reference Bureau, unpublished data, 2000). However, in the DHS data, observations are retrospective over 40 years, and there is an indication that infant mortality occurred at a higher rate in earlier years, although it was not statistically significant. Of those who died, 53.4 percent were male, and 77.6 percent were living in rural parts of the country. Infant mortality was higher for children of mothers younger than 20 years (21.4 percent) and for first-born children (20.8 percent). Infant mortality was higher (17.5 percent) for children of mothers with no formal education than for those whose mothers had secondary or higher education (8.0 percent).

Three spatial Bayesian models were fitted. A baseline model (model 0) included no covariates but an overall constant and site-specific random effects. Model 1 was an extension of the baseline model with the inclusion of year of birth and socioeconomic and demographic variables as potential risk factors. Model 2 included the same parameters as did model 1 but, in addition, adjusted for levels of malaria endemicity. In addition, a Bayesian nonspatial analog of model 2 was fitted for comparative purposes. The parameter estimates obtained from models 1 and 2 and from the nonspatial model are shown in table 1. Estimates of the odds ratios indicate that, in the nonspatial analysis, infant mortality was related to the estimated malaria prevalence after adjustment for the other risk factors. After taking into account the spatial correlation that was present in the infant mortality risk data, we found that the effect of malaria transmission was no longer significant. In fact, the point estimates of the odds ratios change little, but the confidence intervals become wider, confirming the importance of taking spatial correlation into account when analyzing geographic data (16). Model comparison revealed that the model with the smallest deviance information criterion value and, therefore, with the best fit was spatial model 1, which does not include malaria risk.

The fixed-effects parameters of the best fit model 1 showed well-known patterns and confirmed most of the results obtained from the crude data summaries (table 1). In particular, non-firstborn children are at higher risk than their firstborn siblings (odds ratio (OR) = 1.10, 95 percent confidence interval: 1.01, 1.20 for second- or thirdborn; OR = 1.35, 95 percent confidence interval: 1.17, 1.52 for seventh or higher in the birth order). The discrepancies between model-based estimates and observed frequencies in the estimates of the parameters for birth order and preceding birth interval can be explained by the high correlation of the two variables, which introduces confounding. The model allows adjustment for confounders and provides estimates of the effects of one factor in the presence of the other. Infants born

to mothers with no education are at higher risk than those born to mothers with secondary education or higher (OR = 0.55, 95 percent confidence interval: 0.43, 0.74). Mortality was related to the sex of the infant, with boys being at higher risk of dying during the first year of life than girls (OR = 1.15, 95 percent confidence interval: 1.10, 1.21). On the other hand, longer birth intervals reduce the risk of infant death (OR = 0.40, 95 percent confidence interval: 0.35, 0.47 above 4 years vs. less than 2 years). Infants born to older mothers and in urban areas have higher chances of surviving to their first birthday, and there is an indication that infants born in recent years have better survival chances.

The parameters  $\sigma^2$  and  $\phi$  (table 1) measure the variance of the spatial process and the rate of correlation decay (smoothing parameter), respectively. Our data set indicates a small value of  $\phi$  with a posterior median of 0.04 (95 percent credible interval: 0.007, 0.24), suggesting a strong spatial correlation, because this parameter measures the range of the geographic dependency, which is defined as the minimum distance at which spatial correlation between locations is below 5 percent. In our exponential setting, it can be calculated as  $3/\phi = 75$  degrees of longitude and latitude. This implies a nonvanishing correlation between all sampled points and results in very smooth maps for the predicted random effects.

Figure 1 displays the distribution of malaria surveys. The figure shows that most of the surveys were carried out in the south and southwest of the country. There are very few surveys in the center around the Niger River and no surveys in the north of Mali. Model predictions of malaria risk at the DHS locations (figure 2) indicate a low-malaria-risk zone north of the Niger River that contradicts empirical evidence of high transmission, suggesting that the lack of data leads to imprecise estimates in this part of the country, which may distort the relation between malaria endemicity and infant mortality. The unadjusted map of infant mortality risk (figure 3), obtained by the predictions of the Bayesian model, reveals that the highest infant mortality rates are found in the central and central-east part of the country around the Niger River. Distinct foci of high mortality can be identified in the regions of Nara, Banamba, Dioila, Kadiolo, Kolondieba, and Kenieba. Figure 4 represents the variation in infant mortality that is not explained by socioeconomic differences (on a logit scale). This map is a measure of our estimate of the geographic variation of risk of infant mortality, independent of the particular socioeconomic circumstances of the mother. It therefore reflects the marginal burden of infant mortality that is due to ecologic factors, such as malaria transmission intensity and other diseases, at the map locations. A component of this may be residual socioeconomic factors that our covariate data did not fully take into account. In this map, we can distinguish three zones of high risk: the one in the central and central-east border with Burkina Faso, a zone of south Mali that covers the regions from Nara and Diema to Kolondieba and Bougouni, and a zone in the southwest in the region of Kenieba. Estimates of the variance of the residual spatial variation (figure 5) show lower variance in the estimates near locations with observed mortality.

**TABLE 1. Infant mortality estimates for Mali from Demographic and Health Surveys' data in 1995–1996 in combination with malaria risk extracted from the MARA\*/ARMA\* database**

	Crude data		Nonspatial		Model 1		Model 2	
	No. of births	IMR* per 1,000	OR*	95% CI*	OR	95% CI	OR	95% CI
Year of birth								
1960–1965	299	230.8	1.16	0.89, 1.56	1.16	0.88, 1.56	1.16	0.89, 1.56
1966–1971	1,766	239.0	1.41	1.15, 1.79	1.42	1.09, 1.85	1.42	1.12, 1.79
1972–1977	4,200	186.4	1.14	0.93, 1.41	1.15	0.90, 1.49	1.15	0.91, 1.42
1978–1983	7,709	172.8	1.09	0.92, 1.35	1.08	0.84, 1.40	1.09	0.86, 1.34
1984–1989	11,087	163.2	1.05	0.88, 1.30	1.05	0.81, 1.35	1.06	0.85, 1.30
1990–1996	10,845	152.6	1		1		1	
Residence								
Rural	25,615	183.9	1		1		1	
Urban	10,291	132.1	0.68	0.64, 0.72	0.74	0.69, 0.80	0.73	0.68, 0.80
Mother's education								
Secondary or higher	1,416	79.8	0.54	0.42, 0.73	0.55	0.43, 0.74	0.55	0.42, 0.73
Primary	3,563	149.3	0.95	0.80, 1.11	0.95	0.79, 1.15	0.96	0.80, 1.13
None	30,927	175.4	1		1		1	
Sex								
Female	17,718	159.6	1		1		1	
Male	18,188	178.3	1.15	1.10, 1.20	1.15	1.10, 1.21	1.16	1.10, 1.21
Birth order								
1 (firstborn)	7,680	208.3	1		1		1	
2–3	11,746	158.6	1.10	1.02, 1.17	1.10	1.01, 1.20	1.10	1.01, 1.21
4–6	10,851	154.3	1.17	1.06, 1.29	1.15	1.05, 1.31	1.17	1.06, 1.30
≥7	5,629	165.7	1.34	1.20, 1.52	1.35	1.17, 1.52	1.34	1.17, 1.55
Preceding birth interval								
<2 years	18,149	213.8	1		1		1	
2–4 years	15,231	131.8	0.59	0.56, 0.62	0.58	0.53, 0.63	0.59	0.54, 0.63
>4 years	2,526	88.7	0.40	0.36, 0.44	0.40	0.35, 0.47	0.39	0.32, 0.46
Mother's age at infant's birth								
<20 years	9,070	213.9	1		1		1	
20–29 years	19,087	156.6	0.79	0.75, 0.83	0.78	0.72, 0.83	0.78	0.73, 0.83
30–39 years	7,163	146.2	0.73	0.67, 0.78	0.72	0.65, 0.80	0.72	0.65, 0.80
40–49 years	586	160.4	0.82	0.69, 1.02	0.81	0.66, 1.01	0.83	0.66, 1.02
Malaria endemicity								
0.0–0.15	9,951	157.8	1				1	
0.16–0.35	5,159	185.3	0.96	0.88, 1.04			0.97	0.87, 1.09
0.36–0.64	16,824	170.9	0.88	0.83, 0.94			0.89	0.78, 1.10
0.65–1.0	3,972	168.4	0.85	0.77, 0.92			0.92	0.80, 1.08
$\sigma^2$ †					0.88	0.21, 3.82	0.88	0.20, 3.56
$\phi$ ‡					0.04	0.007, 0.24	0.05	0.008, 0.30
DIC*			31,794.21		31,682.64		31,755.73	

\* MARA, Mapping Malaria Risk in Africa; ARMA, Atlas du Risque de la Malaria en Afrique; IMR, infant mortality rate; OR, odds ratio; CI, confidence interval; DIC, deviance information criterion (a measure of model fit for the comparison of models, with smaller values of DIC indicating superior fit).

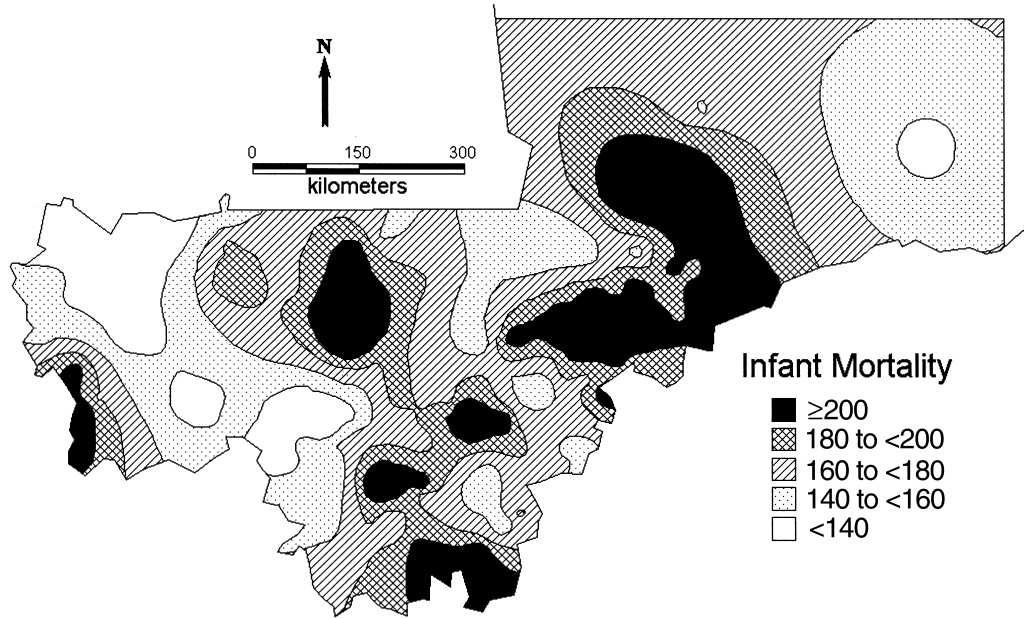
†  $\sigma^2$  is an estimate of the geographic variability.

‡  $\phi$  is the smoothing parameter (see Appendix).

## DISCUSSION

This analysis demonstrates the use of Bayesian geostatistical models in assessing risk factors and producing smooth maps of infant mortality risk from spatially correlated disease data on individuals, such as those available from the

DHS. Results confirmed strong geographic differences in mortality risk and the importance of a number of risk factors such as maternal education and age, birth order and interval, sex, and residence. Year of birth appeared not to be significantly associated with mortality during the first year of life, except for the period 1966–1971, in which a statistically

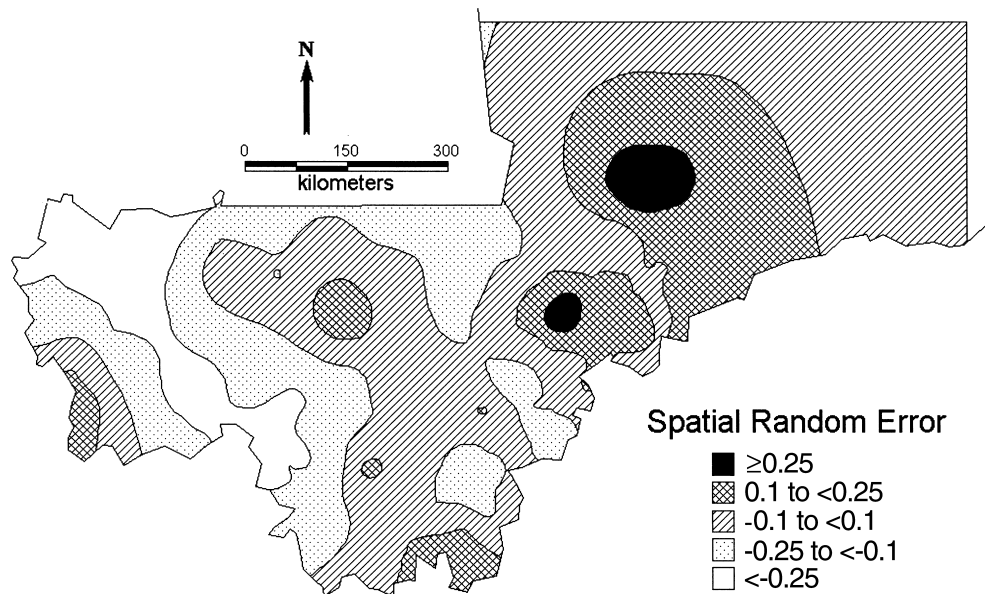


**FIGURE 3.** Smoothed map of the infant mortality rate (per 1,000) in Mali based on the baseline model without covariates.

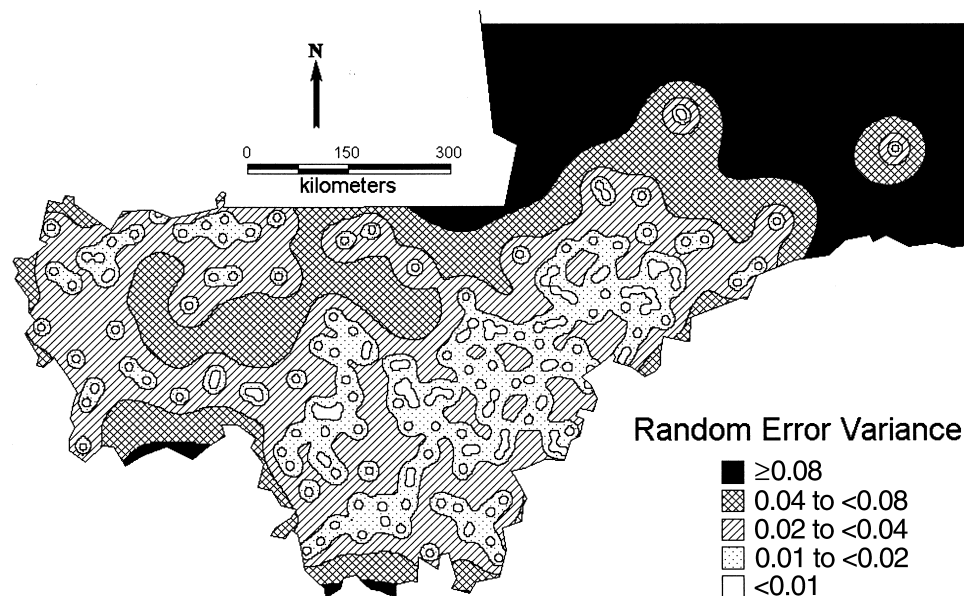
significant increase in infant mortality was observed. There were no differences in infant mortality among the four categories of malaria endemicity defined using the model of Kleinschmidt et al. (11), suggesting that the geographic distribution of malaria is not a major determinant of the pattern of infant mortality in Mali. This finding was not supported by the nonspatial analysis, since accounting for

spatial correlation results in more precise estimates of the standard error and widens the confidence limits of the estimated odds ratios.

The risk factors that were found to be related to mortality are already well known. The negative association between maternal education and mortality has been previously described by Farah and Preston (22). Higher education may



**FIGURE 4.** Map of Mali showing the spatial random effects at the log-odds scale for the socioeconomic-adjusted model.



**FIGURE 5.** Map of Mali showing the variance of the residual spatial variation of the infant mortality risk (at logit scale) adjusted for socio-economic variables.

result in greater health awareness, better utilization of health facilities (23), higher income and the ability to purchase goods and services that improve infants' health (24), longer birth intervals, and possibly higher maternal ages (25).

Although several studies have investigated the relation between malaria transmission and child mortality, this is the first study to our knowledge that attempts to assess this relation by taking into account the geographic variation that is present for both parameters after adjustment for socio-economic confounders. In this analysis, we consider malaria prevalence as a measure of malaria transmission. There are alternative indicators of transmission intensity that have been used to study the effects of malaria on mortality. However, the relation between these indicators has not been fully investigated. The most usual measure is the entomologic inoculation rate, which is the product of the vector-biting rate multiplied by the proportion of mosquitoes infected with sporozoite-stage malaria parasites. Beier et al. (26) report that the entomologic inoculation rate is only weakly related to malaria prevalence. To our knowledge, no studies have been carried out on the best measure of transmission to study mortality.

We have estimated for the first time the geographic distribution of the burden of infant mortality in Mali in addition to that which can be attributed to socioeconomic differences. It is plausible that a large measure of this burden is due to the effect of malaria on infant mortality, even though we were not able to demonstrate this directly. The lack of a relation between malaria risk and infant mortality could reflect unmeasured local factors, for instance, variations in health provisions or the availability of a water supply in the dry Sahel region, which could have a stronger influence than malaria risk on infant mortality. Such unmeasured covariates

with spatial structure could also explain the residual spatial correlation in the data. Information missing from the database regarding malaria control measures taken at different locations could also confound the analysis. Methodological problems related to the compilation of survey data sets, such as the MARA database, constitute a further limitation to our analysis. The surveys, for example, were carried out at different seasons and include different age groups at the various locations. It may well be that, in areas with seasonal malaria, the effect on mortality would be stronger than in areas with perennial malaria. The model of Kleinschmidt et al. (11) for malaria prevalence did not take seasonal variation into account, although malaria transmission in Mali is known to be highly seasonal (F. C. Tanser, South African Medical Research Council, Hlabisa, South Africa, unpublished manuscript).

An additional problem with the database is the sparsity of the surveys in the central-east part of the country. In this analysis, we used a subset of the data in order to deal with the different age groupings of the surveys at the various locations. Currently we are working on alternative approaches that overcome the limitations of heterogeneous age grouping, without omitting data.

An additional methodological problem is the misalignment of the DHS and MARA surveys in time and space. Our analysis is based on the assumption that spatial patterns of infant mortality and malaria risk are relatively stable in time. Although this assumption can be questioned, statistical analysis of the temporal changes of malaria prevalence over the last 40 years at the country level (27) showed no significant patterns. In addition, our analysis of infant mortality rates indicates a statistically significant time trend only for the early years of 1966–1971, but it is quite stable for the last 20

years, when over 80 percent of all cases are recorded. To overcome geographic misalignment, we estimated malaria prevalence at the DHS locations using the spatial malaria model of Kleinschmidt et al. (11). Although we believe that this modeling approach gives a good estimate of the general pattern of malaria prevalence in Mali and of the overall populations at risk, we cannot be confident in local malaria predictions, especially in areas remote from sampled locations. In particular, the paucity of sampling points in areas of very high infant mortality, especially in the Niger River delta, may have resulted in poor predictions in these areas. We propose to address this problem by compiling the databases from a larger area of West Africa and analyzing only the data points where misalignment is minimal. Despite these limitations, our study has demonstrated the considerable potential of spatial statistical methods for analyzing the DHS data. To our knowledge, this is the first analysis of infant mortality to use geostatistical models. The methods presented are valuable for both producing smoothed (covariate-adjusted) maps of mortality risk and assessing covariate effects. Such maps are particularly helpful to identify high mortality areas for most efficiently allocating limited resources in child survival programs.

## ACKNOWLEDGMENTS

The work of the first author was supported by Swiss National Foundation grant 3200-057165.99.

The authors are grateful to the MARA collaboration for making the Mali malaria data available and to Macro International, Inc., for providing the names of the villages of the DHS.

## REFERENCES

- Payne D, Grab B, Fontaine RE, et al. Impact of control measures on malaria transmission and general mortality. *Bull World Health Organ* 1976;54:369–77.
- Bradley DJ. Malaria. In: Feachem RG, Jamison DT, eds. *Disease and mortality in Sub-Saharan Africa*. New York, NY: Oxford University Press, 1991:190–202.
- Snow RW, Marsh K. New insights into the epidemiology of malaria relevant to disease control. *Br Med Bull* 1998;54:293–309.
- Molineaux L. La lutte contre les maladies parasitaires: le problème du paludisme, notamment en Afrique. (In French). In: Vallin J, Lopez A, eds. *La lutte contre la mort*. Paris, France: Presses Universitaires de France, 1985:111–40. (Travaux et documents no. 108).
- Smith TA, Leuenberger R, Lengeler C. Child mortality and malaria transmission intensity in Africa. *Trends Parasitol* 2001;17:145–9.
- Snow RW, Marsh K. Will reducing *Plasmodium falciparum* transmission alter malaria mortality among African children? *Parasitol Today* 1995;11:188–90.
- Steketee RW, Nahlen BL, Parise ME, et al. The burden of malaria in pregnancy in malaria-endemic areas. *Am J Trop Med Hyg* 2001;64(suppl):28–35.
- Mapping Malaria Risk in Africa/Atlas du Risque de la Malaria en Afrique (MARA/ARMA). Towards an atlas of malaria risk in Africa. First technical report of the MARA/ARMA collaboration. Durban, South Africa: MARA/ARMA, 1998. (<http://www.mara.org.za/>).
- Rip MR, Keen CS, Kibel MA. A medical geography of perinatal mortality in metropolitan Cape Town. *S Afr Med J* 1986;27:399–403.
- Kalipeni E. Determinants of infant mortality in Malawi: a spatial perspective. *Soc Sci Med* 1993;37:183–98.
- Kleinschmidt I, Bagayoko M, Clarke GPY, et al. A spatial statistical approach to malaria mapping. *Int J Epidemiol* 2000;29:355–61.
- Hutchinson MF, Nix HA, McMahon JP, et al. Climate data: a topographic and climate database. (CD-ROM). Canberra, Australia: Centre for Resource and Environmental Studies, the Australian National University, 1995.
- World Resources Institute. Africa Data Sampler: a geo-referenced database for all African countries. (CD-ROM and user's guide). Washington, DC: World Resources Institute, 1995.
- GDE Systems, Inc. Geoname Digital Gazetteer. Version I. (CD-ROM). San Diego, CA: GDE Systems, Inc, 1995.
- Coulibaly S, Dicko F, Traoré SM, et al. Enquête démographique et de santé Mali 1995–1996. (In French). Calverton, MD: Celule de Planification et de Statistique du Ministère de la Santé, Direction Nationale de la Statistique et de l'Informatique et Macro International, Inc, 1996.
- Cressie NAC. *Statistics for spatial data*. New York, NY: Wiley, 1993.
- Diggle PJ, Tawn JA, Moyeed RA. Model-based geostatistics. *Appl Stat* 1998;47:299–350.
- Gelfand AE, Smith AFM. Sampling-based approaches to calculating marginal densities. *J Am Stat Assoc* 1990;85:398–409.
- Smith AFM, Roberts GO. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *J R Stat Soc (B)* 1993;55:3–23.
- Gelfand AE, Ravishanker N, Ecker M. Modelling and inference for point-referenced binary spatial data. In: Dey D, Ghosh S, Mallick B, eds. *Generalized linear models: a Bayesian perspective*. Basel, Switzerland: Marcel Dekker, Inc, 1999:373–86.
- Spiegelhalter DJ, Best N, Carlin BP, et al. Bayesian measures of model complexity and fit. *J R Stat Soc (B)* 2002;64:583–639.
- Farah AA, Preston SH. Child mortality differentials in Sudan. *Popul Dev Rev* 1982;8:365–83.
- Jain A. Determinants of regional variation in infant mortality in rural India. In: Jain A, Visaria L, eds. *Infant mortality in India: differentials and determinants*. New Delhi, India: Sage Publications, 1988:127–67.
- Schultz TP. Interpretation of relations among mortality, economics of the household, and the health environment. In: *Proceedings of the Meeting on Socioeconomic Determinants and Consequences of Mortality*, Mexico City, June 19–25. Geneva, Switzerland: World Health Organization, 1979.
- Cleland J, van Ginneken JK. Maternal education and child survival in developing countries: the search for pathways of influence. *Soc Sci Med* 1989;27:1357–68.
- Beier JC, Killeen GF, Githure JI. Short report: entomologic inoculation rates and *Plasmodium falciparum* malaria prevalence in Africa. *Am J Trop Med Hyg* 1999;61:109–13.
- Snow RW, Omumbo JA, Lowe B, et al. Relation between severe malaria morbidity in children and level of *Plasmodium falciparum* transmission in Africa. *Lancet* 1997;349:1650–4.



## APPENDIX

Let  $Y_{ij}$  be a binary response corresponding to the survival status of child  $i$  at site  $s_j$ ,  $j = 1, \dots, n$  taking value 1 if the child is alive after the first year and 0 otherwise, and let  $\mathbf{X}_{ij}$  be the vector of associated covariates. Following the modeling framework of Diggle et al. (17), we introduce the unobserved spatial variation by assuming a latent stationary Gaussian process  $\mathbf{U}(s)$  over our study region,  $\mathcal{D}$ , such that  $\mathbf{U} = (U(s_1), U(s_2), \dots, U(s_n)) \sim N(0, \Sigma)$ , where  $\Sigma_{ij}$  is a parametric function of the separation vector  $s_i - s_j$ . Conditional on  $\mathbf{U}$  and the regression coefficients  $\boldsymbol{\beta}$ , the  $Y_{ij}$  are independent Bernoulli variates with survival probabilities  $p_{ij}$  given by  $\text{logit}(p_{ij}) = \mathbf{X}_{ij}\boldsymbol{\beta} + U(s_j)$ . We assume an isotropic spatial process with  $\Sigma_{ij} = \sigma^2 \rho(s_i - s_j; \phi)$  and an exponential correlation function  $\rho(s_i - s_j; \phi) = \exp(-\phi d_{ij})$ , where  $d_{ij}$  measures the Euclidean distance between the sites  $s_i$  and  $s_j$ .

To complete the Bayesian model specification, we need to adopt prior distributions for the model parameters. We chose noninformative uniform priors for the regression coefficients, that is,  $\boldsymbol{\beta} \propto \mathbf{1}$ , and the following vague priors for the  $\sigma^2$  and  $\phi$  parameters:  $\sigma^2 \sim \text{inverse gamma}(a_1, b_1)$  and  $\phi \sim \text{inverse gamma}(a_2, b_2)$ , with  $a_1 = a_2 = 2.01$ ,  $b_1 = b_2 = 1.01$ . The model was fitted using Markov chain Monte Carlo simulation and, in particular, Gibbs sampling (18). The posterior distribution of  $\sigma^2$  then is conjugate inverse gamma. From the nonstandard one-dimensional conditional distributions of all components of  $\boldsymbol{\beta}$ ,  $\mathbf{U}$  such as  $\phi$ , we sampled by using a random walk metropolis algorithm having a Gaussian proposal density with the mean equal to the estimate from the previous iteration and the variance derived from the inverse second derivative of the log-posterior. We ran a single chain sampler with a burn-in of 5,000 iterations, with convergence assessed by inspection of ergodic averages of selected model parameters. The chain thereafter sampled every 60th iteration until a sample size of 2,000 had been attained.

For model comparison, we utilize the deviance information criterion (DIC), as recently proposed by Spiegelhalter et al. (21). For a vector of parameters  $\boldsymbol{\theta}$ , it is defined by  $\text{DIC} = 2\bar{D} - D(\bar{\boldsymbol{\theta}})$ , with  $D(\cdot)$  being the deviance statistic  $D(\boldsymbol{\theta}) = -2 \log p(\mathbf{y}|\boldsymbol{\theta}) + 2 \log f(\mathbf{y})$ .  $\bar{D}$  is the posterior expectation of  $D$ , and  $\bar{\boldsymbol{\theta}}$  is the posterior expectation of  $\boldsymbol{\theta}$ , with both of them easily estimated from outputs of the Markov chain Monte Carlo simulation. Smaller values of the deviance information criterion indicate better fitting models.

To produce a smooth map of mortality risk, we use Bayesian kriging (16, 20). In particular, we obtain estimates of the mortality risk,  $\mathbf{Y}_0 = (Y(s_{01}), Y(s_{02}), \dots, Y(s_{0l}))$  at any unsampled location  $s_0 = (s_{01}, s_{02}, \dots, s_{0l})$  by the predictive distribution

$$P(\mathbf{Y}_0|\mathbf{Y}) = \int P(\mathbf{Y}_0|\boldsymbol{\beta}, \mathbf{U}_0)P(\mathbf{U}_0|\mathbf{U}, \sigma^2, \phi)P(\boldsymbol{\beta}, \mathbf{U}, \sigma^2, \phi|\mathbf{Y}) d\boldsymbol{\beta} d\mathbf{U}_0 d\mathbf{U} d\sigma^2 d\phi, \quad (1)$$

where the distribution of  $\mathbf{U}_0$  at new sites given  $\mathbf{U}$  at observed sites is normal

$$P(\mathbf{U}_0|\mathbf{U}, \sigma^2, \phi) = N(\Sigma_{01}\Sigma_{11}^{-1}\mathbf{U}, \Sigma_{00} - \Sigma_{01}\Sigma_{11}^{-1}\Sigma_{10}), \quad (2)$$

with  $\Sigma_{11} = E(\mathbf{U}\mathbf{U}^t)$ ,  $\Sigma_{00} = E(\mathbf{U}_0\mathbf{U}_0^t)$ ,  $\Sigma_{01} = \Sigma_{10}^t = E(\mathbf{U}_0\mathbf{U}^t)$  and  $p(Y(s_{0i})|\boldsymbol{\beta}, U(s_{0i})) \sim \text{ber}(p(s_{0i}))$ , with  $\text{logit}(p(s_{0i})) = \mathbf{x}(s_{0i})\boldsymbol{\beta} + U(s_{0i})$ . Equation 1 is the expectation  $E[P(\mathbf{Y}_0|\boldsymbol{\beta}, \mathbf{U}_0)P(\mathbf{U}_0|\mathbf{U}, \sigma^2, \phi)]$  over the posterior distribution  $P(\boldsymbol{\beta}, \mathbf{U}, \sigma^2, \phi|\mathbf{Y})$ , which is identified by the Gibbs sampler. Numerically, this expectation is approximated by the average

$$\frac{1}{r} \sum_{k=1}^r P(\mathbf{Y}_0^{(k)}|\boldsymbol{\beta}^{(k)}, \mathbf{U}_0^{(k)})P(\mathbf{U}_0^{(k)}|\mathbf{U}^{(k)}, \sigma^{2(k)}, \phi^{(k)}), \quad (3)$$

where  $(\boldsymbol{\beta}^{(k)}, \mathbf{U}^{(k)}, \sigma^{2(k)}, \phi^{(k)})$  are samples drawn from the posterior  $P(\boldsymbol{\beta}, \mathbf{U}, \sigma^2, \phi|\mathbf{Y})$ . For mapping purposes, predictions were made for 600,000 pixels covering a regular grid of the whole area of Mali south of 18 degrees latitude north.