The effect of endemicity on the sensitivity and specificity of malaria case definition and attributable fraction in northeast Tanzania

M.L. KAMUGISHA

Amani Medical Research Centre, P.O. Box 4, Amani, Tanga, Tanzania

e-mail: mkamugisha@hotmail.com

Abstract: Data obtained from cross-sectional surveys conducted in Muheza, Tanzania was used to estimate the proportion of fever cases in the population, which could be attributable to malaria. For comparative purposes, two approaches were used to estimate the attributable fraction in lowland and highland areas. The fraction of fever cases at each level of parasite density was used to estimate the sensitivity and specificity of alternative malaria case definitions. Using classical methods, the overall attributable fraction was 29.1% in the lowland areas compared to 33.8% in the highland areas. The model-based method indicated an attributable fraction of 13.2% and 17.3% in the lowland and highland areas, respectively. For the lowland areas, the sensitivity was found to be 100% and the specificity 21% for fever cases (a body temperature $\geq 37.5^{0}$C) with parasitaemia. In the highlands, by contrast, the sensitivity was 100% and the specificity 43%. Both areas provided a similar parasite density cut-off point at the optimum threshold. Based on these findings, it can be concluded that the level of endemicity may not have much impact on the sensitivity of clinical malaria case definitions.

Key words: malaria, case definition, endemicity, attributable fraction, Tanzania

Introduction

Most clinical episodes of malaria comprise of a febrile illness with rather non-specific symptoms. In highly endemic areas in Africa most of the people, especially children, have malaria parasites but are not at that time suffering from malaria disease (Schellenberg *et al.*, 1994). This shows how important it is to distinguish between the disease caused by malaria parasites and the frequent asymptomatic infection caused by the same parasites. In most health facilities in Tanzania, microscopy is the most widely used laboratory-based diagnostic test for malaria. The Tanzanian National Malaria Diagnosis and Treatment Guidelines for clinical malaria defines those patients with a history of fever or a body temperature of $\geq 37.5^{0}$C as being malaria cases (MoH, 2001). However, in most cases, a febrile child is usually diagnosed as a malaria case without examination of blood slide to confirm the presence of malaria parasites.

Although clinical malaria usually present with fever, general body malaise and chills, these signs are not malaria specific and may be due to other infections, which are presented with similar signs (Luxemburger *et al.*, 1998). Presumptive treatment is a strategy for malaria management in most rural health facilities in endemic countries where laboratory facilities are poor. The potential benefits of presumptive treatment, however, have not been adequately measured in areas with low malaria transmission. For instance, in areas with low malaria transmission in north-east Tanzania presumptive treatment of malaria has been found to be unreliable due to high chances of misdiagnoses (Massaga *et al.*, 1999), and that the level of misdiagnosis increases with age and is likely to be significantly higher during the dry season, when

transmission is relatively low. Presumptive treatment based on fever is therefore, less effective and may result into a large proportion of children being inappropriately treated as malaria cases.

The study reported in this paper aimed at assessing the case definitions for clinical malaria for areas with different levels of endemicity by using the attributable fraction approach obtained from modelling the relationship between fever risk and parasite density (Schellenberg *et al.*, 1994; Smith *et al.*, 1994). Specifically, the study aimed at (i) comparing the estimated malaria-attributable fever in the highland and lowland areas of north-eastern Tanzania; (ii) comparing the malaria-attributable fractions estimated by classical methods with model-based approaches; and (iii) estimating the sensitivity and specificity at different parasite density cut-off points.

Materials and methods

Study area

Data used in this analysis were obtained from a study conducted in Muheza district, north-eastern Tanzania in 1998 (courtesy of Prof. C.F. Curtis). The district lies between 4°45'S and 39°00'E. Topographically, it is diverse, with 3 distinct geographical zones (lowland, intermediate and highland) ranging from 200-1200 m above sea level. It has a tropical climate with two main rainy seasons (long rains between March and May, and short rains between October and November). The annual rainfall averages 1800 mm in the highlands and 1000 mm in the lowlands, and is spread over a greater number of days. The ambient temperature ranges from 20-30°C and the relative humidity 65-100%. The district has a population of approximately

400,000, of whom 20% are children aged less than 10 years. Most of the inhabitants are subsistence farmers.

In the district, *Plasmodium falciparum* is the major malaria parasite accounting for about 90% of all infections. *P. malariae* and *P. ovale* are also present, but at a lower incidence (Fowler *et al.*, 1993).

Study design

The district was divided into two zones, i.e., lowlands (<900m) and highlands (>900m). A sample of 11 and 8 villages was randomly selected from the highland and lowland zones, respectively. At the beginning of the study a census of the population was carried out and personal details were recorded. The villages involved were mapped, each house was allocated a project number and children aged 1-14 years were recruited and given a unique identification number after the oral consent had been obtained from either parents or guardians.

A monthly cross-sectional survey was conducted in the 19 selected villages. For each survey a random sample of 50 children drawn from the census were requested to attend, and for each child, date of birth, history of fever for the previous week, and use of antimalarial drugs were recorded. Body temperature was measured, and spleen size estimated. Finger prick blood samples were obtained for preparation of thick films for malaria diagnosis. For each child the body temperature was taken (using a Thermoscan thermometer) and recorded twice from one ear and once from the other. The arithmetic mean temperature was calculated in order to obtain the average body temperature for each child.

Data analysis

Data was analysed using STATA statistical package version 5 (Statistic Corporation, Texas, United States). A fever case was defined as having (on average) a body temperature $\geq 37.5^{0}C$. A reported fever case was defined as having a history of fever for the past one week. Those who were reported as having a history of fever by their mothers or guardians were also regarded as fever cases. In this context attributable fraction (AF) was defined as the proportion of fever cases above the cut-off points attributable to malaria. The population attributable fraction (PAF) was defined as the proportion of all fever cases attributable to malaria. The calculation of the fraction of fever cases at each level of parasite density was used to estimate the sensitivity and specificity of alternative malaria case definitions.

Setting malaria case definitions using different accepted parasite thresholds was used to differentiate between individuals whose fever was caused by *P. falciparum* from those due to other causes. Sensitivity and specificity were used to validate malaria case definition. Sensitivity was defined as the total number of fever due to malaria with a parasite density above the selected cut-off divided by the total number of malaria attributable fever cases with any malaria parasites. Specificity was defined as the total number of fever cases without malaria, below the parasite density cut-off over the total number of fever cases without malaria parasite.

Since this data is from a cross-sectional study, the prevalence of the disease was used to estimate the fraction of prevalent febrile episodes that are attributable to the presence of malaria in the blood slide. The population attributable fraction was calculated from single binary exposure (Bruzzi *et al.*, 1985; Smith *et al.*, 1994).

$$PAF = (pf - pa)(1 - pa)$$

Where: PAF = population attributable fraction
pf = prevalence of parasites among febrile
pa = prevalence of parasite among afebrile

Sensitivity and specificity values at various density cut-off limits were derived from 2x2 contingency tables of the diagnosis results against disease status in a cross-sectional study.

Sensitivity = estimated number of fever cases identified correctly by the case definition divided by the total number of fever cases attributable to malaria.

Sensitivity = $(n_k * Af_k)/(N*PAF)$

Where

n_k = *number of fever cases with malaria above parasite density cut-off*
N = *estimated total number of fever cases*

Specificity = estimated number of fever cases without malaria given cut-off point identified correctly by the case definition divided by the total number of fever cases not attributable to malaria.

Specificity = $(N*(1 - PAF) - n_k *(1 - AF_k)/(N*(1 - PAF)$

The variance of estimated attributable fraction (AF) was obtained by applying the delta method (Benichou, 1991) as:

$$Var(AF) = mn_o(n_1 m_o m + nn_o m_1)/n^3 m^3_o$$

Where:
n_1 = *number of exposed fever cases*
n_o = *number of unexposed fever cases*
n = *total number of exposed and unexposed fever cases*

m_1 = *number of exposed controls*
m_0 = *number of unexposed controls*
m= *total number of exposed and unexposed fever cases*

The corresponding 100(1- a) per cent confidence interval (CI) for attributable fraction (AF) is given by: AF - $z_{1-a/2}$*Var(AF)^0.5

The data was stratified by area of residence (highland/lowland) and the logistic regression was used to model fever as a continuous function of the parasite density. In this model the relationship between parasite density and malarial fever was modelled as defined by Smith *et al.* (1994):

$$\log it = \alpha + \beta * Pd^\tau.$$

Where: *Pd = parasite density*

The probability that an individual case of fever was attributable to malaria was estimated by calculating first the odds ratio for each individual using the logistic model and secondly for each model fitted as described below (Bruzzi *et al.*, 1985).

The attributable fraction: $AF = \dfrac{1}{n}\sum_{i=1}^{n}(\dfrac{r_i-1}{r_i})$

Where:

    r_i = *estimated odds for the ith individual*
    n= number of febrile cases

The probability that any individual fever case is malaria-attributable was obtained from the above logistic regression model. Then probability was used to estimate the proportion of diagnosed fever cases that were attributable to malaria at each case definition, using a parasite density cut-off. The proportion of these diagnosed cases that are attributable to malaria was estimated by:

$$AF_c = \dfrac{1}{n_c}\sum_{i=1}^{n} d_{c,i}\dfrac{(r_i-1)}{r_i}$$

Where:

    AF_c= *The proportion of the diagnosed fever cases that are attributable to malaria*
    c =   *Parasite density cut-off point*
    n_c =   *fever cases with densities in excess of c*

$d_{c,i} = \begin{cases} 1 \\ 0 \end{cases}$ *1 if fever cases satisfy the malaria case definition, 0 otherwise.*

Therefore, the number of fever cases identified correctly by case definition was estimated as the product of the proportion of the diagnosed fever cases

that were attributable to malaria and fever cases with parasite density in excess of the cut-off *c*.

The total number of malaria-attributable fever cases was obtained by multiplying the estimated attributable fraction from logistic regression by the total number of fever cases (PAF).

The parasite density as the linear function was compared with the LR statistic from power model. Tsiatis (1980) method for the goodness-of-fit was carried out to test the adequacy of fit of the model. The odds ratio obtained from model were used to estimate the attributable fraction and to estimate the sensitivity and specificity profiles for a range of malaria case definition thresholds by using the attributable fraction estimated from the model. Selection of the expected confounding variables to be included in the model was done by using stepwise maximum likelihood estimation at significant level of $P = 0.05$. The model was adjusted for potential expected confounding variables that are age, haemoglobin, surveys, sex and weight. None of these variables was included in the logistic regression model at the significant level of $P = 0.05$ in both study areas.

Results

A total of 5425 children participated in the cross-sectional surveys. Table 1 and 2 represents the summary for the variables used in the analyses. Forty-seven percent (2542) of the children were from the lowlands and 53% (2883) were from the highlands. There was no significant difference between sex ratios in the two areas of residence ($P = 0.1$). A total of 1285 (51%) and 1256 (49%) of the children in lowlands were females and males, respectively; whereas from the highlands, 1456 (50.5%) and 1427 (49.5%) were females and males respectively.

The estimated odds ratio for being a fever case comparing parasitaemia and aparasitaemia for the two areas were 1.41 (95% CI: 103 – 1.94, *P*-value = 0.03) and 1.51 (95% CI: 1.05 – 2.18, *P*-value = 0.02) for lowlands and highlands, respectively. The estimated attributable fractions were 23.9% and 21.8% of all fevers attributable to malaria in the lowlands and the highlands, respectively (Table 1). Combining all data gives an odds ratio of 1.77 and an attributable fraction of 33.2%. The above estimated attributable fraction obtained from odds ratio of being a fever case (comparing parasitaemia and aparasitaemia), provides the proportion of fever, which would be prevented from the population if malaria cases were eliminated.

Table 1: Attributable fraction for different cut-offs points in the lowlands and highlands

State of fever	Parasite density					
	Any parasite		< 500 parasites/μl		> 5000 parasites/μl	
	Low	High	Low	High	Low	High
Febrile	232	84	61	1408	58	30
Afebrile	1666	1408	1012	769	230	163
Odds ratio	1.41	1.51	1.23	2.09	2.19	4.43
AF	0.291	0.338	0.19	0.521	0.54	0.774
PAF (%)	23.9	21.8	9.6	24.5	11.1	17.9

The estimate of attributable fraction differed in different selected parasite cut-off points. Sensitivity and specificity profiles were estimated for a range of malaria case definition thresholds using the estimates obtained from the attributable fraction. The intersection between the sensitivity and specificity curves was used as the optimum threshold for sensitivity and specificity profiles.

The parasite density was re-categorised into seven parasite groups, with approximately the same number of observations in each category. The frequency and cumulative distribution in each group is shown in Table 2. These groups were used as the cut-off points for the parasite density (malaria case definition thresholds). The odds ratio of being a fever case by parasite density cut-off point was calculated using the classical method. Because of the nature of discrete variables it was not easy to estimate the

to 68.1% and specificity from 63.6 to 74.5% at the parasite density cut-off point of around 881 – 4560 parasite/μl. For highlands some of the odd ratios estimated were less than one. The estimated value of the odds ratio less than 1 gave a negative attributable fraction, although, the prevalence of the febrile cases was greater than the afebrile cases. The highland data was re-categorised into five groups of parasite density as 0, 1-499, 5000-49999 and 50000+. Still an estimate of odds ratio less than one, which resulted into negative attributable fraction, was obtained.

The estimated attributable fraction for the linear model was less than that estimated by the power model in the lowlands and highlands, respectively (Table 2). The estimated attributable fraction for the highlands was 17.3% and for the lowlands 13.2%.

Table 2: Estimates of attributable fraction for linear regression and power regression

Area	Type of model	Estimated coefficient (β)	LR statistic	Deviance	AF	Goodness-of-fit statistic (P-value)
Highlands	Linear model	0.0000448	48.49	497.07	0.167	2.91 (0.41)
	Power model	0.0000668	49.39	496.6	0.173	2.63 (0.45)
	τ	0.965				
Lowland	Linear model	0.0000386	32.04	866.1	0.121	2.36 (0.8)
	Power model	0.0000826	32.50	865.9	0.132	2.18 (0.82)
	τ	0.93				

optimum thresholds for sensitivity and specificity profiles at a specific parasite density. However, the sensitivity and specificity of the optimum thresholds were estimated based on the range of parasite density groups.

The results from lowland areas showed that the optimum thresholds for sensitivity ranged from 64.7

The estimated sensitivity and specificity from the power model are shown in Table 3. The optimum thresholds (point of intersection) were 3540 parasites/μl in the lowlands and highlands, respectively. These points had sensitivity and specificity of 84% for both areas.

Table 3: Sensitivity and specificity estimated using attributable fraction

Parasite cut-off point	Lowlands		Highlands	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
Optimum threshold	84	84	89	89
Any parasite	100	21	100	43
500/ml-5000/µl	97	55	99	64
>5000/µl	79	87	88	90

Discussion

Two methods, namely, the classical and model-based approaches were used for estimating the attributable fraction and sensitivity and specificity of selected parasite density cut-off points in lowlands and highlands of Muheza district in Tanzania. At the cut-off point of a temperature $\geq 37.5^0$C together with any asexual *P. falciparum*, the classical method shows an over-estimation of the attributable fraction for malaria cases for fever episodes in both areas. However, modelling the parasite density as the linear function gives an under-estimate of the attributable fraction for both areas as compared with the power model.

These findings indicate the crucial importance of the definition of the baseline when estimating the attributable fraction. Benichou (1991) pointed out that the influence of the baseline definition on the attributable fraction can be appropriately interpreted relative to the definition of the baseline. In this study it was observed that the results using classical methods depended on the categories formed.

The sensitivity obtained from the lowlands showed that the optimum threshold for sensitivity ranges from 64.7 to 68.1%, and that the specificity ranges between 63.6% and 74.5% at a parasite density cut-off point of around 881-4560 parasites/µl. Nonetheless, in the highlands the presence of negative attributable fraction for some of the parasite density group meant that sensitivity and specificity could not be obtained due to lack of monotonicity in the observed relationship between fever cases and parasite density in the highlands data.

The attributable fraction, obtained using power estimation, seemed to be higher in the highlands than the lowlands, which is contrary to the estimation obtained by using classical approach which showed that the lowland areas had a higher attributable fraction when compared to the highlands. The reason is likely to be due to the high proportion of malaria febrile cases, which was observed in the lowlands.

The observed differences in the cut-off points obtained by the two approaches may be due to the binary factors that give different fractions with a different categorization of variables when the classical methods are used (Benichou, 1991). According to Benichou (1991) the use of classical methods cannot be adjusted

for several covariate factors. On the other hand, the model-based approach for estimating the attributable fraction has the advantage of adjusting several covariate variables (confounding factors) used (Benichou, 1991).

The attributable fractions estimated in this study are different from those obtained in Ghana, Gambia and Kilombero, Tanzania. This may be due to differences in the study design and data collection techniques. For comparative purposes, there is need of conducting respective studies using a standard protocol. The findings that progeny thresholds depend on endemicity has also been observed in the two areas with different levels of endemicity as reported by Trappe *et al.* (1985). In another study, Smith *et al.* (1994a) showed that the relationship between parasite density and fever is age dependent. This may be the case in our study, although analysis to confirm this was not carried out. Based on these findings, it can be concluded that the level of endemicity may not have much impact on the sensitivity of clinical malaria case definitions.

Acknowledgements

References

Benichou, J. (1991) Methods of adjustment for estimating the attributable risk in case-control studies: a review. *Statistics in Medicine* **10**, 1753-1773.

Bruzzi, P., Green, S.B., Byar, D.P., Brinton, L.A. & Schairer, C. (1985) Estimating the population attributable risk for multiple factors using case control data. *American Journal of Epidemiology* **122**, 904-915.

Fowler, V.G., Lemnge, M.M., Irare, S.G., Malecela, E., Mhina, J., Mtui, S., Mashaka, M. & Mtoi, R. (1993) Efficacy of chloroquine on *Plasmodium falciparum* transmitted at Amani, eastern Usambara Mountains, north-east Tanzania. *Journal of Tropical Medicine and Hygiene* **96**, 337-345.

Luxemburger, C., Nosten, F., Kyle, D.E., Kiricharoen, L., Chongsuphajaisiddhi, T. & White, N.J. (1998) Clinical features cannot predict a diagnosis of malaria or differentiate the infecting species in children living in area of low transmission. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **92**, 45-49.

Massaga, J.J., Kamugisha, M.L. & Salum, F.M. (1999) Malaria presumptive treatment in Tanzania: is it a rational approach for malaria management in rural health units in an area with low transmission? *African Journal of Health Sciences* **6**, 22-26.

MoH (2001) *National Guidelines for Malaria Diagnosis and Treatment in Tanzania Mainland.* Ministry of Health, Dar es Salaam, Tanzania.

Schellenberg, J.A., Smith, T., Alonso, P.L. & Hayes, R. (1994) What is clinical malaria? Finding case definitions for field research in highly endemic areas. *Parasitology Today* **10**, 439-442.

Smith, T., Schellenberg, J.A. & Hayes, R. (1994a) Attributable fraction estimates and case definitions for malaria in endemic areas. *Statistics in Medicine* **13**, 2345-2358.

Smith, T., Genton, B., Bafa, K., Gibson, N., Taime, J., Naraba, A., al-Yaman, F., Beck, H.P., Hii, J. & Alpers, M.P. (1994b) Relationship between *Plasmodium falciparum* infection and morbidity in a highly endemic area. *Parasitology,* **109**, 539-549.

Trape, J.F., Peelman, P. & Marault-Peelman, B. (1985) Criteria for diagnosing clinical malaria among a semi-immune exposed to intense and perennial transmission. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **79**, 435-442.

Tsiatis, A. (1980) A note on goodness-of-fit tests for logistic regression model. *Biometrica* **67**, *250-251.*