LAUTECH Journal of Engineering and Technology 16 (2) 2022: 158-165

ESTIMATION OF ATTRIBUTABLE FRACTION FOR MALARIA IN PREGNANT WOMEN

[1]Efuwape, B. T., [1]Abdullah, K-K. A., [1]Olasupo, A. O., [1]Adesina, A. O., [2]Efuwape T. O., [1]Abulele I. and [3]Aideyan, D. O.

[1]Department of Mathematical Sciences, Olabisi Onabanjo University, Ago-Iwoye
[2]Bells University, Sango – ota.
[3]Taraba State University
*Corresponding Author. Email: efuwapebt@yahoo.com

ABSTRACT

This study applied the estimation of logit model of fever risk as a disease function of parasite density, age and season to give a more precise estimate. The data used for this study was obtained from the University College Hospital (UCH). Asymptomatic carriage of malaria parasites occurs frequently in endemic areas and the detection of parasites in a blood film from a febrile individual which does not necessarily indicate clinical malaria. In areas of very high transmission such estimates of the attributable fraction may be imprecise because very few individual pregnant women are without parasites. Furthermore, non-malaria fevers appear to suppress low levels of parasitaemia resulting in biased estimates of attributable fraction. We therefore, propose a qualitative response regression model for obtaining precise estimates of the probabilities of pregnant women with different level of parasitaemia having fever due to malaria. Logistic regression methods which model fever risk as a continuous function of parasite density, age, and season to give a more precise estimates than simple analyses of parasite prevalence and overcome problems of bias caused by the effects of non-malaria fevers. The result indicate that age is not a predicting factor affecting the pregnant women living in endemic areas, and also season has a slight effect while the parasite level is a major factor.

Keywords: *Malaria, Parasitaemia, Fever, Logit Regression, Attributable fraction*

INTRODUCTION

Malaria in pregnancy is an obstetric, social and medical problem requiring multidisciplinary and multidimensional solution. Pregnant women constitute the main adult risk group for malaria and 80% of deaths due to malaria in Africa occur in pregnant women. In Africa, 30 million women living in malaria-endemic areas become pregnant each year. For these women, malaria is a threat both to themselves and to their babies, with up to 200, 000 newborn deaths each year as a result of malaria in pregnancy (WHO, 2004). Pregnant women are particularly vulnerable to malaria as pregnancy reduces a woman's immunity to malaria making her more susceptible to malaria infection.

Generally, Malaria can be diagnosed by the presence of parasites and high temperature (fever) that is temperature greater than or equal to 37.5°C. In endemic areas, pregnant women can tolerate parasite without symptoms and may have fever due to other causes. Non-malaria appears to suppress low level parasitaemia resulting in biased estimates of the attributable fractions.

A simple statistical approach to estimate attributable fraction is using the frequency of malaria among fever cases to be compared to the prevalence p_f in febrile pregnant women who can be malaria or non-malaria cases, with the parasites prevalence p_a in afebrile controls from the community.

$$AF = \frac{p_f - p_a}{1 - p_a} \quad \text{(Greenwood 1987)}$$

Where AF = Attributable fraction

p_a = parasite prevalence among those without fever.

p_f= parasite prevalence (irrespective of parasite density) among febrile individual.

However, in area of high transmission (endemic area), p_a is very high and the proportion of afebrile pregnant women without parasitaemia is low, resulting either in negative estimates of AF (when $p_f$is less than p_a) or imprecise estimates when p_a is close to 1.

Aim of the Study

This study investigates the estimation of logit model of fever risk as a disease function of parasite density, age, and season to give a more precise estimate than simple analyses of parasite prevalence and overcome problems of bias caused by the effect of non-malaria fevers.

1. To determine attributable fraction estimates of malaria in pregnant women.
2. To determine factors associated with presence of fever.

Review of Related Literatures

Greenwood *et al* (2007) and Rougemont *et al* (2010) have previously presented a method of attributable fraction estimates for areas in West Africa with seasonal malaria transmission by comparing the parasite prevalence, P_f in febrile children who can be malaria or non-malaria cases with the parasite prevalence, P_a in afebrile controls using the classical estimate of the population attributable to a given exposure and is given by

$$AF = p\,(R-1)/R$$

Where p is the proportion of cases exposed, and R is the relative risk of disease associated with the exposure. In this context, the cases are fever episodes and exposure is the presence of malaria parasites in the blood slide. Taking $p=P_f$ and estimating R by the odds ratio, $P_f(1-P_a)/P_a(1-P_f)$, after simplification they obtain:

$$AF = \frac{p_f - p_a}{1 - p_a}$$

This method can be applied to data either from cross-sectional surveys, in which some observations (the 'cases') are in febrile children, the remainder acting as "controls" or cases control studies in which fever cases are ascertained at a health facility and compared with a suitable chosen sample of febrile controls.

In areas of very high endemicity, where parasite prevalence in young children may exceed 80%, this approach is of limited value. The precision of the estimate depends heavily on the proportion of afebrile children without parasitaemia, and when this is small the attributable fraction estimate has very wide confidence limit. Moreover, it is known that other febrile illnesses, including measles and influenza, may suppress malaria parasiteamia (Rogier, et al., 2009). Thus, the parasite prevalence in non-malaria fever cases may be lower than in comparable healthy controls, causing bias in the estimate of *AF*. High temperatures inhibit the development of P. *falciparum* in human beings, so that asymptomatic low-density malaria infections may become sub-patent when children develop fevers from other causes.

Smith *et al* (2010) showed that in some cases the method proposed by Greenwood *et al* (2011) is not precise. The method did not work well for the Kilombero study because of the small number of aparasitaemia samples. In fact, the method breaks down completely in the study as they compare the distribution of parasites densities when children were febrile with observations on afebrile children. The overall parasites prevalence in febrile episodes was lower than that in the afebrile controls, giving a negative estimate of the attributable fraction.

To overcome these problems, Smith *et al* (2007) explored alternative methods which make use of the quantitative data on parasite density. Finally, they presented alternative approaches to attributable

fraction estimation and illustrate them using data from highly endemic rural community in Tanzania (Kilombero District Morogoro Region). Method for attributable fraction estimates which include logistic regession models fitted for kilombero data such as model for effect of parasite density and model for effects of confounders (method of age adjustment). This method was able to overcome the problem of negative of imprecise estimates (bias).

Vounatsou, *et al* (2000) was able to come up with another interesting method which is the most recent for attributable fraction estimates by proposing a novel approach for obtaining precise estimates of the frequency of clinical malaria by formulating the problem as a mixture of distributions. The mixture consists of parasite densities in children with fever either due to malaria or due to other causes. The mixing proportion estimates the proportion of children whose fever is attributable to malaria.

The development of simulation-based methods in Bayesian statistics (Smith and Roberts, 1993) has made possible the implementation of the Bayesian approach to mixture problems. Diebolt and Robert (1994) and Robert (1996) derived the posterior distribution using conjugate priors and introduced indicator variables to classify observations, according to the component of the mixture from which they come. Robert (1996) proposed reparametrization of the posterior. It remains unclear how this can be applied in a completely nonparametric approach.

METHODOLOGY

For estimation purposes, we consider

$$L_i = \ln(P_i/(1-P_i)) = B_1 + B_2X_i + U_i \quad . \qquad (1)$$

In order to estimate (1) we need, apart from X*i*, the values of the regressand or logit, L*i*. This depends on the type of data we have for analysis. Distinguish two types of data.

- Data at the individual level
- Grouped or Replicated data

Data at the Individual Level

If we have data on individuals ordinary least square estimation of (1) is infeasible. This is easy to see. In terms of the data used in this study

P*i* = 1, if an individual have malaria parasites accompany with fever and

P*i* = 0, if it does not have fever.

But if we put these values directly into the logit L*i* we obtain

L*i* = ln (1/0) if a pregnant woman have malaria parasites accompany with fever

L*i* = ln (0/1) if a pregnant woman have malaria parasites but no fever

Obviously, these expressions are meaningless, therefore if we have data at the micro, or individual level, we cannot estimate (1) by the standard ordinary least square (OLS) routine. In this situation, we may have to resort to the maximum likelihood (ML) method to estimate the parameters.

Considering the data given in this study, letting Y= 1 if an individual (pregnant woman with fever) have a malaria parasite accompany with fever and Y= 0 if there is some parasite presents but no fever as the prediction. The logit model can be written as:

$$L_i = \frac{Pi}{1-Pi} = B_1 + B_2 \text{ Season} + B_3 \text{ Age} + B_4 \text{ No of parasites} + U_i \ldots \qquad (2)$$

Here, neither ordinary least square (OLS) nor weighted least square (WLS) is helpful. To solve this problem, we have to resort to nonlinear estimating procedure using the method of maximum likelihood. Since most modern statistical packages have routines to estimate logit models on the basis of ungrouped data and interprets the results using Minitab or Eview.

Grouped or Replicated Data

Supposed we have some pregnant women grouped or replicated (repeat observation) according to number of parasites and the number of pregnant

women having fever at each level of number of parasites, corresponding to each number of parasites X_i, there are N_i women, n_i among whom have fever $(n_i \leq N_i)$ therefore, if we compute.

$$\hat{P} = \frac{n_i}{N_i} \quad . \quad (3)$$

that is, the relative frequency, we can use it as an estimate of the true P_i corresponding to each X_i, if Ni is fairly large, $\hat{P}_i$ will be a reasonably good estimate of P_i.

Recall from elementary statistics, that the probability of an event is the limit of the relative frequency as the sample size becomes infinitely large. Using the estimated P_i, we can obtain the estimated logit as

$$\hat{L}_i = In\left(\frac{\hat{P}_i}{1-\hat{P}_i}\right) = \hat{B}_1 + \hat{B}_2 X_i. \quad . \quad (4)$$

It can be shown that Ni is fairly large and if each observation in a given range of number of parasites X_i is distributed independently as a binomial variable.

$$\mu_i \sim N\left[0, \frac{1}{N_i P_i(1-P_i)}\right] \quad (5)$$

.

that is U_i follows the normal distribution with zero mean and variance equal $1/[N_i P_i(1-P_i)]$ because the disturbance term is the logit model is heteroscedastic, instead of using OLS we will have to use the weighted least squares (WLS). For empirical purposes, however, we will replace the unknown P_i by $\hat{P}_i$ and use

$$\hat{\sigma}^2 = \frac{1}{N_i \hat{P}_i\left(1-\hat{P}_i\right)} \quad (6)$$

as estimator of σ^2

Estimation of Grouped Logit Regression

In this section, we present the various step for estimating a grouped logit model. For each range of number parasites X, we compute the probability of having fever as $\hat{P}_i = n_i/N_i$

For each X_i, obtained the logit as

$$\hat{L}_i = In\left[\hat{P}_i \Big/ \left(1-\hat{P}_i\right)\right]$$

To resolve the problem of heteroscedasticity, transform (1) as follows.

$$\sqrt{w_i}L_i = B_1\sqrt{w_i} + B_2\sqrt{w_i}X_i + \sqrt{w_i}U_i \quad (7)$$

which can be written as

$$L_i^* = B_{1i}\sqrt{w_i} + B_2 X_i^* + v_i \quad (8)$$

Where the weights $w_i = N_i\hat{P}_i\left(1-\hat{P}_i\right)$;

L_i^* = transformed or weighted L_i

X_i^* = transformed or weighted X_i

v_i = transformed error term.

It is easy to verify that the transformed error term v_i is homoscedastic, keeping in mind that the original error variance is $\sigma_u^2 = 1/[NiPi(1 - Pi)]$

To estimate (7) by OLS – recall that weighted least square (WLS) is OLS on the transformed data. Notice that in (7) there is no intercept term introduced explicitly (Why?). Therefore, one will have to use regression through the origin routine to estimate (7).

Set up confidence intervals and/or test hypotheses in the usual OLS framework, but keep in mind that all the conclusion will be valid strictly speaking if the sample is reasonably large.

ANALYSIS OF DATA AND INTERPRETATION

Analysis of Data on Individual Pregnant Woman

Considering the data used in this project, letting Y=1 if a pregnant woman has parasites accompany with fever. i.e temperature greater than or equals 37.5°C and Y=0 if a pregnant woman has malaria parasites but no fever. Age and season are also used as the malaria predictor. The logit model can be written as:

$$Li = \frac{Pi}{1-Pi} = B_1 + B_2 season + B_3\, age + B_4 no\ of\ parasites + U_i \quad (9)$$

TABLE 1: THE LOGISTIC REGRESSION RESULT				
Dependent Variable: Temperature				
Method: ML - Binary Logit (Newton-Raphson)				
Convergence achieved after 5 iterations McFadden R-square= 0.210				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-2.373151	0.613088	-3.870818	0.0001
SEASON	-0.070172	0.330790	-0.212136	0.8320
NAGE	-0.281470	0.160679	-1.751754	0.0798
NPARASIT	0.827642	0.101202	8.178135	0.0000
Prob (LR statistic)	0.000000			
Obs with Dep=0 304		Total obs=393		
Obs with Dep=1 89				

Since we are using the method of maximum likelihood, which is generally a large sample method, the estimated standard errors are *asymptotic*.

As a result, instead of using the t-statistics to evaluate the statistical significance of a coefficient, we use the (standard normal) Z statistics. So inferences are based on the normal table. Recall that if the sample size is reasonably large, the t distribution converges to the normal distribution.

The conventional measure of goodness of fit, R^2 is not particularly meaningful in binary regression model measure similar to R^2, called pseudo R^2 are available and there are variety of them.

Eviews presents one such measure, the Mc Fadden R^2, denoted by R^2 McF

Like R^2, R^2 McF also ranges between 0 and 1

Another comparatively simple measure of goodness of fit is the count R^2 which is define as.

$$\text{Count } R^2 = \frac{number\ of\ correct\ predictions}{total\ number\ of\ obsevations}$$

It should be noted, that in binary regression model, goodness of fit is of secondary importance. What matters is the expected signs of the regression coefficient and their statistical or practical significance.

To test the null hypothesis that all the slope coefficients are simultaneously equal to zero, the equivalent of F test in the linear regression model is the likelihood ratio (LR) statistic.

Given the null hypothesis, the LR statistic follows the chi-square distribution with degree of freedom equal to the number of explanatory variables.

Interpretation of the Result

Each slope coefficient in the equation is a partial slope coefficient and measures the change in the estimated logit for a unit change in the values of a given regressor (holding other regressors constant). Thus, the NPARASITES coefficient of 0.828 means, with other variables held constant that if NPARASITES increase by a unit, on the average, the estimated logit increases by about 0.828 unit suggesting a positive relationship between the two.

A more meaningful interpretation is in terms of odds which are obtained by antilog of the various slope coefficients. Take the antilog NPARASITES i.e (2.287) $e^{0.827642}$. This suggests that pregnant women who have a particular number of parasites are more than two times likely to have fever due malaria than those who have lesser parasites and other things remaining the same.

Suppose we want to compute the actual probability of a pregnant woman having malaria, we put the actual data of the pregnant woman in the estimated logit model, this result in the estimated logit value for the pregnant woman.

Using,

$$P_i = E(Y = 1/X_i) = 1/1 + e^{-(B_1+B_2X_i)}$$

Putting the actual data of the pregnant woman in the estimated logit model given in table 4.1.1.

We arrive at the estimated individual probability of the pregnant woman of having malaria.

Recall the count R^2 defined earlier, the table below gives the actual and the predicted values of the regress and. Since the regress and in the logit model takes a values of 1 or 0, if predicted probability is greater than 0.5, we classify that as 1, but if it is less than 0.5 we classify that as 0, we then count the number of the corrected predictions and compute the R^2 .

COMPUTING PROBABILITY

(ATTRIBUTABLE FRACTION)

Suppose we want to compute probability at X =2723, plugging this value

$$\hat{L}i^* = -2.321708\sqrt{wi} + 0.0001379\hat{X}i^* =$$

$$-2.321708\ (2.7848) + 0.0001379(7583.01)$$

$$\hat{L}i^* = -5.419795$$

And dividing this by $\sqrt{w_i} = 2.7848$ (see Table 1), we obtain $\hat{L}\mathrm{i} = -1.9462$

That is,

$$\hat{L} * / \sqrt{w} = \hat{L}_i = +x$$

Therefore, at the parasite level of 2723, we have

$$-1.94621 = ln\left(\frac{\hat{P}i}{1-\hat{P}i}\right)$$

Therefore, $\left(\frac{\hat{P}i}{1-\hat{P}i}\right) = e^{-1.94621} = 0.1428$(odds ratio)

Solving this for,

$$\hat{P} = \frac{e^{-1.94621}}{1+e^{-1.94621}} = 0.1250$$

That is, given the parasite level 2723, the probability of a pregnant women having malaria is about 13%.

CONCLUSION

The use of glogit regression offers advantages like first, smoothing over parasite densities, the method makes full use of the detailed information on the parasite density distributions in fever cases and control, secondly, adjustment for covariates can be achieved more flexible using ungrouped logit model. The method of attributable fraction estimate lends itself naturally to the estimation of probabilities that each pregnant woman can be attributed to malaria. This study is considered to be important and is seen as a way to develop educational programs that aim at the reduction malaria in pregnant women in our highly endemic environment.

RECOMMENDATIONS

1. Delivery of malaria intervention through antenatal clinics and primary health centre in Nigeria needs to be widespread. This approach is currently the exception rather than the rule. However, a few have already adopted the strategy as policy.
2. As resistance to antimalarial drug increases, the challenges of treatment and prevention of malaria among pregnant women become greater. Research in this area is therefore a high priority. There is also a need for research to develop prevention strategies for women residing in areas of low or unstable transmission.
3. Pregnant women who decided not to attend antenatal clinics or who attend only for the first visit or too late during pregnancy need to be reached out. New strategies should be required to encourage these women to attend antenatal care early and consistently.
4. Increase awareness of problem among communities most affected by malaria so as to avoid the problem caused by malaria.

5. Government should provide pregnant women with Insecticides Treated Net (ITNs) as early in pregnancy as possible and their use should be encouraged for throughout pregnancy and during post-partum period.
6. Finally, within the Roll Back Malaria Global Partnership, WHO should work effectively with governmental, non-governmental bilateral and donor agencies to overcome challenges, meet the Abuja goal and reduce the burden of malaria in pregnancy. The availability of insecticide treated nets, effective intermitted preventive treatment and a means of delivery through antenatal clinics, provide a unique opportunity that must be taken to protect the millions of Nigerians women who become pregnant each year and their babies.

REFERENCES

Agresti, A. (2006). Introduction to categorical Data Analysis. John Wiley and Sons, Inc.

Armstrong S., Smith, T., Alonso, P.L. and Hayes, R.J. (2004). What is clinical malaria? Finding cases definition for field research in highly endemic areas. *Parasitol Today* **10,** 439-442.

Benichou, J. (2001). 'Methods of adjustment for estimating the attributable risk in case-control studies: a review' Statistics in Medicine, **10,** 1753-1773.

D'Alessandro U. (2006). The impact of a national impregnated bed net programme on the outcome of pregnancy in primigravidae in The Gambia. Transactions of the *Royal Society of Tropical Medicine and Hygiene*, 90(5):487-492.

Delley V, Bouvier. P, Breslow. N, Doumbo O, Sagara I, Diakite M, Mauris A, Dolo A, Rougemont A (2000). What does a single determination of malaria parasite density mean? A longitudinal survey in Mali. *Trop Med Int. Health* **5**:404-412.

Diebolt, J. and Robert, C. P. (2004) Estimation of finite mixture distributions through Bayesian sampling.

Goodman CA, Coleman PG, Mills AJ. The cost-effectiveness of antenatal malaria prevention in sub-Saharan Africa. *American Journal of Tropical Medicine and Hygiene*, 2001, 64(1,2 S):45-56.

Goodman C.A., Coleman PG, Mills AJ. (2009). Cost-effectiveness of malaria control in sub-Saharan Africa. *Lancet,* 354:378-385.

Gujarati, D. (2004) Basic Econometrics, McGraw Publication, London,

Greenwood, B. M. (2007) Asymptomatic malaria infections- Do they matter*? Parasitol. Today,* **3(7),** 206- 214.

Kitua, A.Y., Smith, T., Alonso, P. L.,Masanja, H., Menendez, C., Kimario, J and Tanner, M(2006) *Plasmodium falciparum* malaria in the first year of life year in an area of intense and perennial transmission. *Trop. Med. Int. Health,* **1 (4)**, 475-484.

Kwiatkowski, D. (2009). 'Febrile temperatures can synchronize the growth of Plasmodium falciparum in vitro' Journal of Experimental Medicine, 169, 357-361.

Kilombero C.V. (2002). Malaria Project 'The level of anti-sporozoite antibodies in highly endemic malaria area and its relationship with exposure to mosquito' *Transactions of the Royal Society of Tropical Medicine and Hygiene,* **86,**499-504.

Kuile T. (2003) Permethrin-treated bednets reduce malaria in pregnancy in an area of intense perennial malaria transmission in western Kenya. *American Journal of Tropical Medicine and Hygiene*.

Parise M.E. (2008) Efficacy of sulfadoxine-pyrimethamine for prevention of placental malaria in an area of Kenya with a high prevalence of malaria and human immunodeficiency virus infection. *American Journal of Tropical Medicine and Hygiene,* 59(5):813-822.

Rogier C, Commenges D, Trape JF, (2006). Evidence for an age-dependent pyrogenic threshold of *Plasmodium falciparum* parasitemia in highly endemic populations *Am J Trop Med Hygiene* **54**: 613-619.

Rougemont A, Breslow N, Brenner E, Moret AL, Dumbo O Dolo A, Soula G, Perrin L, (2001). Epidemiology basis for clinical diagnosis of childhood malaria in endemic

zone in West Africa village. *Lancet* 338: 1292-1295.

Rothman, K. (2006). *Modern Epidemiology,* Little, Brown and Company, Boston.

Rooth, I. B. and Bjorkman, A (2002). 'Suppression of *Plasmodium falciparum* infections during concomitant measles of influenza but not during pertussis' *American Journal of Tropical Medicine and Hygiene* **47,** 675-681 (2002).

Smith T, Schellenberg J.A, Hayes R. (2004) Attributable fraction estimates and case definitions for malaria in endemic areas *Statis Med*.,**13,** 2345-235

Steketee RW *et al*. The burden of malaria in pregnancy in malaria-endemic areas. *American Journal of Tropical Medicine and Hygiene,* 2001, 64(1,2 S):28-35.

Shulman C. E. (2009). Intermittent sulphadoxine-pyrimethamine to prevent severe anaemia secondary to malaria in pregnancy: a randomised placebo-controlled trial. *Lancet,* 353:632-636.

Schultz L. J. (2004). The efficacy of antimalarial regimens containing sulfadoxine-pyrimethamine and/or chloroquine in preventing peripheral and placental Plasmodium falciparum infection among pregnant women in Malawi. American *Journal of Tropical Medicine and Hygiene*, **51**(5):515-522.

Smith, T. Charlwood, J. D., Kihonda, J., Mwankusye, S., Billingsley, P., Meuwissen, J., Lyimo, E., Takken, W., Teuscher, T. and Tanner, M. (2000). 'Absense of seasonal transmission', Acta Tropica, **54,** 55-72.

Teuscher, T. (2002). 'Household-based malaria control in a highly endemic area of Africa(Tanzania): Determinants of transmission and disease and indicators for Kilombero malaria project' Memorias do Instituto Oswaldo Cruz, **87,** Suppl, 111, 121-130.

Tsiatis, A. (2000). 'A note of goodness-of-fit tests for the logistic regression model' *Biometrika,***67**,1, 250-151.

Wolfe E *et al*. (2001). Cost-effectiveness of sulfadoxine-pyrimethamine for the prevention of malaria-associated low birth weight? *American Journal of Tropical Medicine and Hygiene,* 64(3,4):178-186.

Wiseman V *et al*. (2003). The cost-effectiveness of permethrin-treated bednets in an area of intense malaria transmission in western Kenya. *American Journal of Tropical Medicine and Hygiene*, [in press].