Clinical Case Definitions and Malaria Vaccine Efficacy

William O. Rogers,1,2 Frank Atuguba,2 Abraham R. Oduro,2 Abraham Hodgson,2 and Kwadwo A. Koram1

1Noguchi Memorial Institute for Medical Research, University of Ghana, Legon, and 2Navrongo Health Research Center, Navrongo, Ghana; 3Naval Medical Research Unit No. 3, Cairo, Egypt

Reported efficacies from vaccine trials may depend heavily on the clinical case definition used in the trial. The dependence may be particularly striking for diseases such as malaria, in which no single case definition is appropriate. We used logistic regression modeling of the relationship between parasitemia and fever in data sets from Ghanaian children to determine the fraction of fevers attributable to malaria and to model how the choice of a threshold parasitemia in the clinical case definition affects the measured efficacy of malaria vaccines. Calculated clinical attack rates varied 10-fold as a function of the threshold parasitemia. Strikingly, measured vaccine efficacies in reducing clinical malaria depended heavily on the threshold parasitemia, varying between 20% and 80% as the threshold varied between 1 and 5000 parasites/μL. We suggest that clinical case definitions of malaria that incorporate a threshold parasitemia are arbitrary and do not yield stable estimates of vaccine trial end points.

Vaccine efficacies are calculated by comparing the incidence of disease end points in vaccinated and non-vaccinated subjects. Lack of accuracy in the diagnostic methods used to identify disease and the use of imperfect clinical case definitions may introduce important errors into estimates of vaccine efficacy. The resultant errors are complex and may lead to over- or underestimates of the efficacy. Relatively few attempts have been made to estimate the magnitude of such effects. Blackwelder et al. [1] found that efficacy estimates for acellular pertussis vaccines varied substantially as the clinical case definition was narrowed. Lachenbruch [2] modeled the effect of diagnostic errors on measured efficacies of a hypothetical Lyme disease vaccine and found that significant underestimates of vaccine efficacy resulted when diagnostic specificity fell below 90%. The magnitude of the effect increased markedly as the disease incidence decreased.

The choice of clinical case definitions and end points in malaria vaccine trials is complex and may have marked effects on reported vaccine efficacies. For trials of preerythrocytic-stage vaccines (such as RTS,S [3]), which are designed to prevent any parasitemia, the clinical case definition is straightforward, but diagnostic errors may significantly distort reported efficacies. Ohrt et al. have shown that even a low rate of false-positive diagnoses can lead to substantial underestimates of the efficacy of malaria chemoprophylaxis [4]. In the case of erythrocytic-stage vaccines, which are designed to reduce morbidity by blocking parasite toxins or by reducing parasitemia without necessarily eliminating it entirely, the choice of an appropriate case definition is difficult. In areas of high malaria endemicity, children may tolerate parasitemia without developing clinical illness; indeed, the majority of asymptomatic children may harbor malaria parasites. In such settings, the co-existence of fever and parasitemia does not necessarily indicate clinical malaria, because the parasitemia may be coincidental to fever of another cause. Most clinical trials of malaria vaccines have, therefore, used a clinical case definition that includes fever and a parasitemia that is greater than a chosen threshold. Because high
parasitemias are more likely to cause fever, this is a reasonable approach, but any choice of threshold may have important effects on the estimate of vaccine efficacy.

Smith et al. have proposed a logistic regression modeling approach for the estimation of the fraction of fevers attributable to malaria (\( \lambda \)) and the diagnostic sensitivity and specificity of clinical case definitions that incorporate varying threshold parasitemias [5]. We use this approach to estimate \( \lambda \) in data sets from 2 sites in Ghana. We demonstrate the relative insensitivity of these estimates to a series of potential confounding variables, and we use the estimates of \( \lambda \) to derive the sensitivity, specificity, and positive and negative predictive values of clinical case definitions of malaria based on threshold parasitemias at the 2 sites. Finally, we assess the sensitivity of measured vaccine trial end points and measured vaccine efficacies to variation in the chosen threshold parasitemia.

**SUBJECTS, MATERIALS, AND METHODS**

*Study sites, designs, and subjects.* Data from 3 separate studies at 2 study sites in Ghana were analyzed. Two studies were conducted in the Kassena-Nankana District (KND) of northern Ghana. This region is located in northern Ghana’s Sahelian savannah belt and experiences intense, seasonal malaria transmission. The characteristics of this site have been described extensively elsewhere [6–11]. In a prospective cohort study, between May and December 2003—the rainy, high-transmission season—we measured the incidence densities of fever and parasitemia in 811 children 6–18 months old (at enrollment) by use of either passive (\( n = 604 \)) or active (\( n = 207 \)) follow-up. Details of this study will be published elsewhere. We analyzed 1488 clinical encounters. In 51.3% (764/1488) of the encounters, the subject was febrile (axillary temperature \( \geq 37.5^\circ \mathrm{C} \)). Additional afebrile control subjects were obtained from among asymptomatic community control subjects who had been enrolled in a case-control study of severe malaria in the KND from July 2002 through December 2003. Among the community control subjects, there were 1217 afebrile subjects 6–24 months old; they were used as afebrile control subjects in some analyses, as described in Results.

A separate analysis was performed on data from a study conducted in the Hohoe District (HD), which is located in the forest belt of the Volta region of Ghana. Malaria transmission in the HD is both less seasonal and less intense than in the KND. Children presenting at an outpatient clinic in the HD for any reason were screened for enrollment into a study of the clinical efficacy of a series of antimalarial drugs from December 1998 through December 2000. The screening data included axillary temperatures and the results of a malaria smear. Of the subjects, 31.4% (303/964) were febrile at the time of screening.

In all cases, parasitemia was estimated by examination of Giemsa-stained thick films. The number of asexual-stage parasites/200 leukocytes was counted, and parasitemia was estimated on the basis of a uniform leukocyte density of 8000 leukocytes/\( \mu \)L.

The use of human subjects in these studies was approved by the scientific and ethics review boards of the Noguchi Memorial Institute for Medical Research, the Navrongo Health Research Center, the Ghanaian Ministry of Health, and US Naval Medical Research Unit No. 3. The studies were conducted in accordance with regulations governing the protection of human subjects in medical research.

*Statistical analysis.* Logistic regression modeling of the risk of fever as a function of parasitemia, estimation of \( \lambda \), and derivation of the sensitivity, specificity, and positive and negative predictive values for clinical case definitions that incorporated different threshold parasitemias were performed as described by Smith et al. [5]. To assess the effects that different clinical case definitions have on measured malaria vaccine efficacies, we made the following assumptions: we treated each data set as representing a set of clinical encounters in a cohort of study subjects during the observation period of a vaccine trial, and we considered a positive clinical episode to be one in which there was both fever and parasitemia greater than a specified threshold. For each threshold parasitemia, we summed the number of true-positive and false-positive episodes of clinical malaria, on the basis of \( \lambda \) calculated from the logistic regression model; this yielded the number of episodes of clinical malaria that would be reported in the control group. We simulated the effect of 2 potential vaccines: (1) an antidisease vaccine that does not inhibit parasite replication but that does block the action of parasite toxins and (2) an antiparasite vaccine that inhibits parasite replication by targeting blood-stage antigens. To model the antidisease vaccine, we left the parasitemias at each observation unchanged but reduced the chance that any given level of parasitemia would produce a fever by 50%. To model the antiparasite vaccine, we reduced the parasitemia in each observation by a constant factor and recalculated the probability of fever on the basis of the results of the original logistic regression of fever versus parasitemia. We then computed the sum of true-positive and false-positive episodes of clinical malaria on the basis of each threshold parasitemia. Measured vaccine efficacy was estimated for each threshold parasitemia as \( 1 - ( \text{total events with vaccine/total events with control} ) \).

Statistical analyses were performed using SPSS (version 10.0), Stata/SE (version 8.00; StataCorp), and SigmaStat (version 2.03; Systat Software).

**RESULTS**

We first analyzed data from a cohort study designed to measure the incidence of clinical malaria in children in the KND of northern Ghana. There were 1488 outpatient episodes from
this study for which data on fever and parasitemia were available. The characteristics of these episodes and the patients who experienced them are shown in table 1. The mean parasitemia was 1.11 log parasites/µL (95% confidence interval [CI], 0.94–1.39 log parasites/µL) greater in febrile subjects than in afebrile subjects.

Table 2 shows the results of logistic regression of fever versus parasitemia. In the simplest model, \( \logit(\pi) = \alpha + \beta(x_i) \) (where \( \pi \) is the probability that the \( i \)th episode is febrile, \( x_i \) is the parasitemia in the \( i \)th episode, and \( \alpha, \beta, \) and \( \tau \) are coefficients in the logistic regression model), the best fit was obtained with \( \tau = 0.24 \), and the calculated \( \lambda \) was 61\% (95\% CI, 56\%–65\%). Because all subsequent calculations depended on the estimated value of \( \lambda \), we considered several covariates and possible sources of error in the logistic regression model used to estimate \( \lambda \). First, because clinical episodes identified by active surveillance may be milder than those identified by passive surveillance, we tested a model in which membership in the active surveillance group or the passive surveillance group was included as a categorical variable. Inclusion of this variable slightly improved the fit of the model but did not significantly alter the estimate of \( \lambda \). Second, we tested models in which age was included as a linear term, as a quadratic term, or as a categorical variable (six 3-month age categories); none of these models significantly improved the fit of the model or altered the estimate of \( \lambda \). Third, because the risk of malaria is not uniform throughout the rainy season, we included the month of occurrence of the clinical episodes as a categorical variable. Inclusion of this variable slightly improved the fit of the model but did not significantly alter the estimate of \( \lambda \). Fourth, to account for the nonindependence of serial observations of individual subjects, we used generalized estimating equations that treated multiple observations of the same subject as a matched set. The estimate of \( \lambda \) was not affected. Finally, we considered the source of the afebrile episodes used in the model. It is possible that the afebrile subjects among our subjects with clinical episodes were more likely to be parasitemic than were the asymptomatic children in the community at large. As part of a separate case-control study of severe malaria in the KND, we had data from 1217 afebrile, asymptomatic children 6–24 months old that had been collected during the 2002 and 2003 high-transmission seasons. We repeated the analysis after replacing the original subjects with afebrile episodes with the afebrile community control subjects from 2002 and 2003. The resultant change in the estimate of \( \lambda \) was small. Because none of the potential covariates or sources of error significantly altered the estimated \( \lambda \), we used the simplest model without covariates for the subsequent analysis.

We considered a series of clinical case definitions of malaria. Each definition required the presence of an axillary temperature \( \geq 37.5^\circ\text{C} \) and a parasitemia above a specified threshold. Figure 1A shows the sensitivity, specificity, and positive and negative predictive values of the clinical case definition as a function of the threshold parasitemia in the KND. Among febrile children presenting for evaluation, the estimated prevalence of clinical malaria (or the prior probability that a febrile child has malaria) is 61\%. A positive malaria smear raises this probability to only 70\%, and even a parasitemia >10,000 parasites/µL raises the probability to only 81\%. The negative predictive value at this threshold is 62\%. In this epidemiological setting of high transmission and substantial clinical immunity, a positive malaria smear contributes relatively little to the diagnosis of clinical malaria, regardless of the threshold parasitemia chosen. Figure 1 shows only the effect of the threshold parasitemia on calculated sensitivity and specificity and presumes perfect microscopy. Because the sensitivity of microscopy decreases sharply at low parasitemias [12], the model will overestimate the true sensitivity and negative predictive value at low parasitemia thresholds.

To assess the effect of a varying parasitemia threshold in a setting of less-intense transmission, we performed a similar logistic regression analysis of \( \lambda \) using data from children in the HD, an area of less-intense malaria transmission, compared with the KND. In the HD, \( \lambda \) was 25.3\%; the estimate was not significantly changed by incorporating the previously considered covariates into the model (data not shown). Figure 1B shows the sensitivity, specificity, and predictive values of the case definition as a function of threshold parasitemia. The prior probability of clinical malaria in a febrile child was 25\%. The finding of any parasitemia raises this probability to 41\%, and the finding of a parasitemia >5000 approximately doubles the probability, to 52\%. At this threshold, the negative predictive value is still high, at 97\%.

Table 1. Characteristics of the subjects with malaria and clinical episodes.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Febrile</th>
<th>Afebrile</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean ± SD, months</td>
<td>14.4 ± 3.8</td>
<td>14.4 ± 3.8</td>
<td>14.4 ± 3.8</td>
</tr>
<tr>
<td>Male sex</td>
<td>402 (52.6)</td>
<td>391 (54.0)</td>
<td>793 (53.3)</td>
</tr>
<tr>
<td>Parasitemia, mean ± SD, log parasites/µL</td>
<td>3.48 ± 1.58</td>
<td>2.37 ± 1.80</td>
<td>2.94 ± 1.78</td>
</tr>
<tr>
<td>Surveillance arm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td>313 (41.0)</td>
<td>330 (45.6)</td>
<td>643 (43.2)</td>
</tr>
<tr>
<td>Passive</td>
<td>451 (59.0)</td>
<td>394 (54.4)</td>
<td>845 (56.8)</td>
</tr>
</tbody>
</table>

NOTE. Data are no. (%) of subjects, unless otherwise noted.
In both the KND and the HD, a substantial portion of episodes identified as clinical malaria were, in fact, not. Such misclassification may significantly affect epidemiological or interventional studies in which the frequency of episodes of clinical malaria is an end point. We first consider the simple case of a study measuring the incidence of episodes of clinical malaria. We make the approximation that the 1488 clinical encounters from the KND data were obtained from a single group of 800 subjects over a 6-month observation period. Figure 2 shows the effect of varying the threshold parasitemia on the estimated incidence of clinical episodes. The reference line shows the best estimate of the true incidence—the incidence of fever multiplied by $\lambda$. It is clear that, depending on the chosen threshold parasitemia, the reported incidence may substantially over- or underestimate the true incidence.

We next consider how such over- or underestimation of disease incidence might affect the reported efficacy of a malaria vaccine. As described in Subjects, Materials, and Methods, we modeled the effect of changing parasitemia thresholds on measured vaccine efficacies for an antidisease vaccine and an antiparasite vaccine. Figure 3 shows the relationship between measured vaccine efficacy and parasitemia threshold for the antidisease vaccine and the antiparasite vaccine in the KND data set and the HD data set. For the antidisease vaccine model, varying the threshold parasitemia across a wide range had only a relatively minor effect on the measured efficacy (figure 3A and 3C). On the other hand, the measured efficacy of the antiparasite vaccine was exquisitely sensitive to the threshold parasitemia chosen, varying from 10% to 96% as the chosen threshold varied from any detectable parasitemia to up to 25,000 parasites/µL. Even within the narrower range of threshold parasitemias commonly used in intervention trials (200–2500 parasites/µL), the measured efficacy would vary from 34% to 63% (KND) or from 17% to 52% (HD) (figure 3B and 3C). For both the antidisease vaccine and the antiparasite vaccine, the measured efficacy was greater in the KND data set than in the HD data set (figure 3C). This result is not surprising, because $\lambda$—and, therefore, the proportion of true-positive episodes among the measured episodes—was higher in the KND data set.

In modeling the antiparasite vaccine, we arbitrarily chose a 90% reduction in the level of parasitemia in each observation as a plausible vaccine effect. However, the marked sensitivity of measured efficacy with respect to threshold parasitemia does not depend on this choice. Figure 3D shows a series of curves plotting measured efficacy versus parasitemia threshold, calculated for a range of vaccine effects (from a 20% to 99% reduction in parasitemia) at each observation. The measured efficacy remained greatly sensitive to the chosen threshold parasitemia over a wide range of vaccine effects. Only when the biological effect of the vaccine is very low—for example, a 20% reduction in parasitemia—is the measured efficacy relatively insensitive to the threshold parasitemia.

**DISCUSSION**

We investigated the relationship between the sensitivity and specificity of varying clinical case definitions of malaria and their likely effect on measured vaccine efficacies in areas of high endemicity in Africa. We found that clinical case definitions commonly used in interventional studies produce substantial numbers of false-positive and false-negative results and that these errors can significantly affect the measured efficacies of vaccines or other interventions.

We considered the common clinical case definition of malaria to be fever (axillary temperature $\geq 37.5^\circ C$) accompanied by parasitemia greater than a specified threshold density. We used logistic regression modeling to estimate $\lambda$ and to assess the effects of threshold parasitemias on the sensitivity, specificity, and positive and negative predictive values for clinical case
definitions. No threshold parasitemia produced a diagnostic accuracy that was very satisfactory. In the KND, the probability that a febrile child had malaria was 61% before microscopy, and this probability increased to only 84% when a threshold parasitemia of 25,000 parasites/μL was used. Even in the HD, where λ was substantially lower (25%), a threshold parasitemia of 5000 parasites/μL was required to reach a positive predictive value of 52%. Depending on the intensity of transmission, the cost and risk of therapy for malaria, the cost of diagnosis, and the possibility of clinical follow-up for a child who was not treated because he or she had parasitemia below the diagnostic threshold, it may be rational to treat all fevers empirically for malaria. In settings of less-intense transmission, determination of λ and estimates of diagnostic sensitivity and specificity for a variety of threshold parasitemias may guide the policy-based choice between provision of empirical therapy for all fevers and provision of therapy on the basis of microscopy results. A recent analysis of data from several sites in Kenya found that the sensitivity and specificity of case definitions of malaria depended heavily on the intensity of transmission [13].

We found that the threshold parasitemia used to define an episode of clinical malaria could markedly affect measured vaccine efficacies. In the case of antidisease vaccines, which have been proposed to prevent illness without inhibiting parasite replication (by, e.g., targeting parasite toxins [14]), the effect of varying the threshold parasitemia on the measured efficacy in the model was modest (figure 3A and 3C)—in the KND data set, increasing the threshold from any detectable parasitemia to 100,000 parasites/μL increased the measured efficacy from 35% to only 45%, assuming that the vaccine reduced the number of true malarial fevers by 50%. At very high parasitemias, a greater proportion of fevers are, indeed, due to parasitemia and, thus, are preventable by the vaccine. The disadvantage of setting the parasitemia threshold at this very high density is that the number of events is reduced, increasing the width of the 95% CIs around the efficacy estimate (figure 3A). A purely antidisease vaccine that has no effect on parasite replication is unlikely to be useful. Successful vaccines will require a significant antiparasite component.

In the case of antiparasite vaccines, efficacy was highly dependent on the choice of parasitemia threshold. Such a vaccine may prevent an episode of clinical malaria in 1 of 2 ways. The reduction in parasitemia reduces the probability of fever, and an afebrile observation would not be counted as a clinical ep-

---

**Figure 1.** Diagnostic properties of clinical case definitions of malaria. A, The Kassena-Nankana District. B, The Hohoe District. Shown are the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for clinical case definitions based on an axillary temperature ≥37.5°C and a parasitemia greater than a specified threshold, calculated as described in Subjects, Materials, and Methods.

---

**Figure 2.** Effect of the clinical case definition of malaria on estimates of the incidence of episodes of clinical malaria in the Kassena-Nankana District (KND). Shown is the measured incidence of clinical episodes as a function of the threshold parasitemia used in the clinical case definition. The dotted line indicates the product of the fraction of fevers attributable to malaria and the incidence of fever. The incidences are calculated from the KND data set, include pooled data from active and passive surveillance cohorts, and do not include adjustments for censored time at risk after antimalaria treatment or absence from the study site. More-precise estimates of the incidence in the KND under active and passive surveillance that include these adjustments will be published elsewhere.
Figure 3. Effect of varying clinical case definitions of malaria on measured vaccine efficacies. Panels A and B show the modeled vaccine efficacies (and the corresponding 95% confidence intervals) as a function of the threshold parasitemia included in the clinical case definition, for antidisease vaccine [A] and antiparasite vaccine [B], on the basis of the Kassena-Nankana District (KND) data set, calculated as described in Subjects, Materials, and Methods. Panel C shows the comparison between the vaccine efficacy estimates based on the KND data set and the Hohoe District (HD) data set. Panel D shows the dependence of the measured efficacy of antidisease vaccine on the parasitemia threshold used in the clinical case definition as a function of vaccine potency. The relationship between measured vaccine efficacy and threshold parasitemia for an antiparasite vaccine was calculated as described in Subjects, Materials, and Methods for a series of vaccine potencies corresponding to a reduction in parasitemia at each observation by a factor of 0.2, 0.5, 0.8, 0.9, 0.95, or 0.99.

The potential magnitude of these effects has important implications for the design and interpretation of trials of interventions aimed at reducing morbidity from malaria. We found that case definitions with lower parasitemia thresholds had poor specificity and produced low efficacy estimates, whereas case definitions with higher parasitemia thresholds produced higher estimates. It has frequently been reported that antimalarial interventions and genetic resistance factors have a greater impact on severe disease than on mild disease. In a recent clinical trial in Mozambique [15], the efficacy of the RTS,S vaccine [3] against all clinical episodes of malaria was reported to be 29.9%, whereas the efficacy against severe episodes of malaria was reported to be 57.7%. Although the models presented here are not directly applicable to a preerythrocytic-stage vaccine such as RTS,S—and although it is possible that there was a differ-
ential effect of the RTS,S vaccine on severe disease, compared with that on mild disease—the different efficacies for mild and severe malaria may result from differences in the sensitivity and specificity of case definitions for mild and severe malaria. Validation of the models presented here will depend on analysis of the results of large-scale field trials of blood-stage vaccines.

It is important to have a measure of vaccine efficacy that does not depend on an arbitrarily chosen parasitemia threshold. One approach is to measure $\lambda$ and the incidence of fever separately in vaccinated and control groups. These results can be used to calculate the reduction in episodes of malarial fever due to the vaccine, without use of an arbitrary parasitemia threshold. For example, in the model based on the KND data set, again assuming that the vaccine reduced parasitemia at each observation by 90%, the estimates of $\lambda$ in the control group and the vaccine group would be 61% and 42%, respectively, and the number of febrile events occurring during the observation period would be 743 and 601, respectively. The efficacy, measured as a reduction in the incidence of fevers correctly attributable to malaria, would thus be $1 - (0.42 \times 601)/(0.61 \times 743) = 44\%$. This approach avoids the use of an arbitrary threshold parasitemia. One might object that this approach does not provide a prospectively defined case definition that can be applied to an individual subject to determine whether the vaccine failed in that individual case. Indeed, it measures efficacy on a population basis and does not attempt to determine whether any individual outcome represents a vaccine failure. Nonetheless, because the algorithm that is used to calculate efficacy can be specified prospectively, this approach is no less objective than approaches based on threshold parasitemias and will provide a more accurate estimate of the efficacy of a vaccine against clinical malaria.

**Acknowledgments**

We thank Thomas Anorigiya and Nathan Mensah, for assistance in data collection; Tom Smith and Lee Hall, for helpful suggestions; and Miguel Hernan, Malla Rao, and Stephen Hoffman, for critical review.

**References**


