Sources of Error in the Estimation of Mosquito Infection Rates Used to Assess Risk of Arbovirus Transmission

Dulce M. Bustamante* and Cynthia C. Lord

Florida Medical Entomology Laboratory, Department of Entomology and Nematology,
University of Florida, Vero Beach, Florida

Abstract. Infection rate is an estimate of the prevalence of arbovirus infection in a mosquito population. It is assumed that when infection rate increases, the risk of arbovirus transmission to humans and animals also increases. We examined some of the factors that can invalidate this assumption. First, we used a model to illustrate how the proportion of mosquitoes capable of virus transmission, or infectious, is not a constant fraction of the number of infected mosquitoes. This assumption relies on the proportion of infected mosquitoes in the population. However, there are many factors that could invalidate this assumption. Environmental variation could introduce unpredictability into how many latent mosquitoes survive to become infectious, and the reliability of estimated infection rates can be compromised by biases (discussed below) introduced during mosquito collecting and virus testing.

INTRODUCTION

There are approximately 100 arboviruses that can infect humans and cause considerable morbidity and mortality and another 40 that can infect livestock with important economic impacts. The prevalence of viral infection in field-collected mosquitoes is a common surveillance indicator that is used to assess the risk of transmission of viruses such as West Nile virus (WNV) or western equine encephalomyelitis virus (WEEV) to humans and domestic animals. Other surveillance indicators of risk include sentinel animal seroconversions and animal or human infection data.

When uninfected mosquitoes acquire virus from a vertebrate host, they go through a latent phase known as the extrinsic incubation period (EIP), which is the time that it takes the virus to disseminate into mosquito organs including the salivary glands from which transmission to subsequent vertebrate hosts occurs. The duration of the EIP varies among mosquito species, individual mosquitoes, and virus species, and is modulated by external factors such as temperature. In this study, we refer to those mosquitoes that have the ability to transmit virus as infectious. Uninfected, latent, and infectious mosquitoes can all potentially be found in a sample of mosquitoes collected for arboviral surveillance. Ideally, we would use that sample to estimate the proportion of mosquitoes in the population that are infectious because this measure would be more directly related to risk of arbovirus transmission. This method would involve individual mosquito testing for virus dissemination to the salivary glands or actual transmission (e.g., salivation into a capillary tube). However, there is no practical way to do this with large numbers of field-collected mosquitoes. Thus, the proportion of infected mosquitoes, or infection rate,* is estimated as a substitute (Figure 1).

It is assumed that when the mosquito infection rate increases, the risk of arbovirus transmission also increases. This assumption relies on the proportion of infected mosquitoes being a good approximation of the proportion of infectious mosquitoes. It also requires reliable estimates of the proportion of infected mosquitoes in the population. However, there are many factors that could invalidate this assumption. Environmental variation could introduce unpredictability into how many latent mosquitoes survive to become infectious, and the reliability of estimated infection rates can be compromised by biases (discussed below) introduced during mosquito collecting and virus testing.

Temperature is one of the best studied factors affecting the EIP and the rate of virus dissemination in a mosquito, thus directly influencing the proportion of latent mosquitoes that become infectious. Laboratory studies have shown that approximately 40% of Culex pipiens mosquitoes infected with WNV will develop a disseminated infection after 15 days of incubation at 20°C, but this fraction increases to almost 90% at 30°C. In the case of Cx. pipiens quinquefasciatus and WNV, 33% of infected mosquitoes will develop disseminated infections after 13 days of incubation at 25°C, and 81% will have disseminated infections after 13 days at 30°C. Dissemination is necessary for viral transmission, and some studies report that 90% of Cx. pipiens mosquitoes with disseminated WNV infections transmit the virus by bite.

Another example of temperature modulating transmission is WEEV vectored by Cx. tarsalis. In a laboratory experiment, 30% of mosquitoes were capable of transmitting virus to chickens after 6 days of incubation at 32°C, but the proportion decreased to 20% after 12 days. At 25°C, only 4% of mosquitoes were capable of transmission after 6 days, but the fraction increased to 35% after 12 days. A different experiment found that approximately 70% of the infected Cx. tarsalis transmitted virus orally to capillary tubes seven days post-infection at 30°C, but after 12 days this fraction had decreased to approximately 12%. At 20°C approximately 25% of mosquitoes were transmitting virus after 7 days, but the percentage increased to approximately 60% after 12 days.

Possible explanations for a reduction in transmission over time at 30–32°C include the ability of some Cx. tarsalis females to modulate viral dissemination from the mesenteron to the salivary glands or a decrease in the secretory function of the salivary glands caused by cell pathologic changes. Dissemination responses to changes in temperature are nonlinear for WEEV in Cx. tarsalis. Clearly, temperature is a critical factor that modifies the response of mosquitoes to the virus, and this induced variation reduces our capacity to predict the...
1173MOSQUITO ESTIMATED INFECTION AND TRANSMISSION

Figure 1. Estimation of the proportion of infected mosquitoes and its use to assess risk because direct estimation of the proportion of infected mosquitoes that are infectious is not practical. For this estimation to be useful in the assessment of risk, it is important to understand the relationships between infected and infectious mosquitoes, and between the population infection rate and its estimate. Both relationships are studied here in Models I and II.

relationship between the proportion of infected and infectious mosquitoes.

In the field, mosquitoes endure daily changes in temperature and variations of the microclimatic conditions in their resting sites during non-active periods. This temperature variability will cause a population of mosquitoes to contain infected individuals in different stages in the development of a disseminated infection. Consequently, mosquito samples taken on consecutive days could yield similar estimated infection rates, but the infectiousness rate could be different from one day to the next.

Ideally, to estimate the proportion of infected mosquitoes in a population, we would take a large sample of mosquitoes using an unbiased trapping method such that all individuals are equally likely to be collected. The mosquitoes would be tested individually for the presence of the virus with an assay that can detect the specific virus every time and at any concentration. Those ideal conditions are usually not met and surveillance methods introduce biases that affect the reliability of the estimated infection rates.

Trapping methods often introduce biases because they target a particular physiologic stage of the adult mosquito population (host-seeking females, gravid females, resting adults) or they attract some mosquito species more than others. Sampling techniques will also affect the sample size obtained, both in absolute numbers of mosquitoes and in the proportion of the population. Because not all mosquitoes are equally likely to be captured, sampling introduces a bias that affects the relationship between the actual population infection prevalence and the estimated infection rate. It is important to remember that in some cases a collecting bias towards a particular mosquito stage in the population, such as gravid females, is desirable to increase the chances of detecting virus presence in the mosquito population.

Other biases are introduced by the virus detection process. Individual testing of mosquitoes in most cases is not feasible for logistic and financial reasons. Thus, a cost-effective alternative to individual testing is to aggregate mosquitoes in groups or pools and test those for virus. There are many assays that can be used to detect virus in mosquito pools including plaque assays in Vero cell culture, TaqMan assays (reverse transcription–polymerase chain reaction [RT-PCR]), VecTest, enzyme-linked immunosorbent assays, or immunoassays. These assays differ in their power to discriminate among viruses and in their ability to detect the virus particles or plaque-forming units (PFU) at different concentrations.

The quantity of virus in a mosquito pool affects the outcome of the different assays because each assay has a threshold concentration below which virus usually cannot be detected. The quantity of virus in mosquitoes is affected by time and
temperature. The mean body titer of Cx. tarsalis females infected with WNV held at temperatures between 22°C and 30°C was significantly higher than the titer in those females held at temperatures between 14°C and 18°C. In another study with Cx. tarsalis and WEEV, body titers after incubation at 15°C ranged between 10^{2.4} and 10^{5.7} PFU/body (mean = 10^{6.0} PFU/body) for mosquitoes that did not transmit virus into a capillary tube (latent), and between 10^{8.8} and 10^{9.1} PFU/body (mean = 10^{9.0} PFU/body) for those mosquitoes that did transmit the virus (infectious). At 30°C, the titers ranged between 10^{5.7} and 10^{6.0} PFU/body (mean = 10^{5.9} PFU/body) for latent mosquitoes, and between 10^{11} and 10^{12} PFU/body (mean = 10^{10} PFU/body) for infectious mosquitoes. Statistical tests comparing those means were not reported by these authors. Virus titer distributions for other mosquito-virus systems in the field or in the laboratory are largely unknown. Failure to detect virus can occur when the infected mosquitoes (or mosquito) present in a pool have low virus titers, and the assay used needs a high concentration of virus particles to produce a positive result.

After mosquito pool assaying, the information on the number of mosquitoes collected in the sample, the number of pools that tested positive for virus, and the number of mosquitoes in each individual pool are used to calculate the estimated infection rate. There are various ways to conduct this calculation but the two most commonly reported are the minimum infection rate (MIR) and the maximum likelihood estimator (MLE) of the proportion of infected mosquitoes. The MIR is the ratio of the number of positive pools to the total number of mosquitoes in the sample. It is by definition the minimum infection proportion and it assumes that only one infected individual is present in a positive pool. The MLE is the value of the proportion of infected mosquitoes $P$, that maximizes the likelihood of $n$ pools of size $m$ to be virus positive, where $P$ is the parameter for a binomial distribution. The MIR is considered appropriate to use when infections in the mosquito population are at low levels, but during periods of high transmission it will largely underestimate mosquito infections. Thus, the use of the MLE coupled with variable size pooling is recommended. A field study comparing the MIR and MLE estimated infection rates (pool size = 5) with the infection rates calculated from individual mosquitoes did not find major differences among the values generated. However, this pool size is rather small for arbovirus surveillance and the lack of a difference might not have much relevance for field studies. The MLE has the advantage that there are algorithms available that consider variations in pool size.

Infection rates (e.g., MIR) are frequently reported as a measure of virus activity among mosquitoes. However, studies assessing the associations between mosquito infection rates and human or domestic animal disease incidence are limited. A recent study used a 25-year weekly data set from Massachusetts to show that human cases of infection with eastern equine encephalitis virus were positively associated with the number of infected Culex melanura per trap night and with weekly estimates of mosquito infection with eastern equine encephalitis virus. The study also observed that the MIR reached values ≥ 1.0 at approximately week 31 in years when human cases were reported, contrary to years without human cases. Important thresholds of mosquito infection above which human illness can be expected have been determined for other viruses. For example, in Arizona, an MIR ≥ 4 for WNV from weekly Culex mosquito pools is considered a high level of viral activity and human cases are expected to occur. In California, an MIR ≥ 5 for WNV, also from weekly Culex mosquito pools, is an indicator of increased risk of transmission and could lead to emergency planning or to the declaration of an epidemic if other risk factors have also increased concurrently. Although associations exist and mosquito infection rates are used in surveillance, infection rates have limitations. The lack of detection of infection in mosquitoes is not necessarily an indication of absence of viral activity. Some virus detection methods could detect unviable virus RNA in mosquitoes (using PCR methods), which would bias estimates of infection, and it is not possible to easily separate infected and infectious mosquitoes. We explored some of the sources of error and limitations of infection rates.

As mentioned earlier, it is often assumed that as the estimated infection rate increases, the risk of transmission of arboviruses to humans and animals also increases. Basic questions arise from this assumption: does the proportion of infectious mosquitoes increase with the proportion of infected mosquitoes, and can we obtain reliable estimates of infection that reflect changes in the proportion of infected mosquitoes in the population? In this study, we used a model to examine how incubation temperature, mosquito survival, mosquito species, and virus species, influence the relationship between the proportion of infected and infectious mosquitoes (Model I, Figure 1). Additionally, we used numerical simulations to examine how the process of mosquito sampling, pooling, and virus testing affect the relationship between the proportion of infected mosquitoes in the population and estimated proportions calculated from field samples (Model II, Figure 1). Specifically, the simulations considered the sample size, the pool size, and virus concentration needed in the pool to show a positive test result.

The results of our study are a reminder of why changes in infection rates are not always associated with changes in risk of transmission and that this can be due to biologic and methodologic factors. They also indicate that our estimated infection rates usually underestimate the population prevalence of infection, that they might not be direct indicators of increases in transmission risk, and that they should always be used in conjunction with other indicators and variables when making an evaluation of virus transmission risk. A more integrated approach toward mosquito surveillance using infection rate in conjunction with other indicators of risk is necessary to assess changes in the risk of arbovirus transmission and to determine the actual need for public health alerts.

**METHODS AND RESULTS**

**Model I: Relationship between mosquito infection and infectiousness.** This model follows a cohort of mosquitoes that became infected with a virus during the first blood meal, and calculates the proportion of those infected mosquitoes that become infectious at different times $t$ as a result of virus dissemination. Data on arbovirus dissemination in mosquitoes from two different and well documented mosquito-virus systems were obtained from the literature and were used as parameters in this model: Cx. tarsalis-WEEV and Cx. pipiens quinguefasciatus-WNV. Two constant incubation temperatures of 20°C and 30°C were considered.

Following previous similar models, the proportion of surviving mosquitoes in a cohort that are infectious at time $t$ ($INS_t$), was calculated as $INS_t = p' \times INF \times D_{inf}$, where $p'$ is
the daily survival rate of the mosquitoes, \( p \); is the proportion of mosquitoes surviving after \( t \) days, \( INF \) is the proportion of mosquitoes at time \( t \) that carry the virus, and \( D_{\text{inf}} \) is the percentage of infected mosquitoes that show dissemination after \( t - t_1 \) days of incubation under a particular temperature. Here, \( t_1 \) is the time (in days) of the first and only infectious blood meal. Dissemination was used here as an approximation for infectiousness for the \( Cx. \) pipiens quinquefasciatus-WNV system.\(^5\) Data on WEEV transmission by \( Cx. \) tarsalis in the laboratory were available.\(^9,10\)

A single value for the proportion of infected mosquitoes (\( INF \)) was considered for the present model. The term \( INF \) included latent and infectious mosquitoes as dissemination progressed over time. It was assumed that 1% of the surviving mosquitoes in the cohort became infected with the virus during the first blood meal at time \( t_c \). This assumption is a combination of the proportion of blood meals taken on infectious hosts and the probability of infection after an infectious meal. Infection rates in cohorts of mosquitoes in the laboratory do not tend to decrease with time at fixed temperatures.\(^5,10\) Thus, the proportion of infected mosquitoes was kept constant at 1% of the surviving mosquitoes at subsequent times.

The model followed the proportion of infected and infectious mosquitoes in a cohort during four gonotrophic cycles. A gonotrophic cycle was considered as the time required for a mosquito to take a blood meal, go through oogenesis, and lay eggs. We assumed that the mosquitoes can take a blood meal immediately after oviposition. We calculated \( INS \) at the end of each of five gonotrophic cycles because they marked the times of a potential blood meal, thus, the times of potential transmission of the virus from the mosquito to a vertebrate host. The times \( t_1 \) at which \( INS \) was calculated were \( t_1 \) = time of the first blood meal and the only blood meal when mosquitoes became infected, \( t_{2,4} \) = times of the second, third, fourth, and fifth blood meals. The time of mosquito emergence was \( t_e \).

Estimates of the minimum duration of the gonotrophic cycle for \( Cx. \) tarsalis and \( Cx. \) pipiens quinquefasciatus are shown in Table 1. It has been observed for other mosquito species that increases in temperature reduce the duration of the gonotrophic cycle. In a laboratory study conducted on \( Anopheles albimanus \), the duration of the gonotrophic cycle was significantly reduced from 88.4 hours (range = 81.8–94.9 hours) at 24°C, to 69.1 hours (range = 64.6–73.6 hours) at 30°C.\(^26\) This finding indicates that an increase of 6°C would reduce the gonotrophic cycle of this mosquito by an average of 19.3 hours. The field calculations of longer-lived \( Cx. \) pipiens quinquefasciatus during the winter in Calcutta (Table 1) and the information for \( An. \) albimanus support our assumption that the gonotrophic cycle length should decrease approximately one day from 20°C to 30°C. In this study, we consider a gonotrophic cycle of five days at 20°C and four days at 30°C for both species. Time from emergence to the first blood meal is approximately two days for \( Cx. \) tarsalis and \( Cx. \) pipiens quinquefasciatus.\(^26,30\) A previous modeling study with \( Cx. \) tarsalis used a value of 2.5 days from emergence to the first blood meal to represent variations in their calculations (2–3 days).\(^26\)

We use continuous values of \( t \) and consider that at 30°C blood meals were taken at 2.5 (first), 6.5, 10.5, 14.5, and 18.5 days of age. Adjusting for a longer lifespan at lower temperatures, we consider that at 20°C the first blood meal was taken at 3.5 days of age and the subsequent meals were taken at 8.5, 13.5, 18.5, and 23.5 days of age.

Two daily survival rates were used in the model. A survival rate of \( p = 0.85 \) was used for the 30°C temperature for both mosquito species and \( p = 0.90 \) was used for the 20°C temperature. The possible effects of age and infection status on mosquito survival were not considered. Estimates of survival rates for \( Cx. \) tarsalis and \( Cx. \) pipiens quinquefasciatus are shown in Table 1. In Calcutta, India, higher survival rates were observed for \( Cx. \) pipiens quinquefasciatus during the winter months than in the summer months.\(^33\) Higher survival rates at cooler temperatures have also been observed for \( Cx. \) tarsalis in the laboratory.\(^34\) These observations support the selection of \( p = 0.90 \) for the 20°C temperature.

Rates of dissemination previously reported in the literature for WNV in \( Cx. \) pipiens complex mosquitoes and for WEEV in \( Cx. \) tarsalis studied under constant temperature conditions are shown in Figure 2. Those dissemination rates were the bases for estimating \( D_{\text{inf}} \) at different times in the model. Given that for all mosquitoes the infectious blood meal was the first one, \( t_1 \) was equal to 3.5 days at 20°C, and 2.5 days at 30°C. The \( D_{\text{inf}} \) values for all the time steps are shown in Figure 2.

The model was implemented using Microsoft Excel (Redmond, WA).\(^35\) The mosquito-virus systems selected to illustrate different patterns of dissemination over time are only two among many types in these systems, and simplifications such as constant survival rates, fixed temperatures, and equating dissemination to infectiousness are acknowledged. However, this model illustrates the variation of patterns that can arise in the relationship between infection and infectiousness and therefore has broader implications.

### Results for Model I

The results of this model illustrate differences in viral dissemination in a cohort of mosquitoes at two temperatures (Figure 3). At 20°C, WEEV in \( Cx. \) tarsalis starts disseminating faster than WNV in \( Cx. \) pipiens quinquefasciatus. In the cohort, the highest proportion of

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration of the gonotrophic cycle and daily survival rates obtained from the literature for ( Culex ) tarsalis and ( Cx. ) pipiens quinquefasciatus</td>
</tr>
<tr>
<td>Species</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>( Cx. ) tarsalis</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>( Cx. ) pipiens quinquefasciatus</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
surviving *Cx. tarsalis* mosquitoes that became infectious occurred 10 days after the infectious blood meal ($t = 13.5$ days). After this time, the proportion of infectious mosquitoes decreased because of mosquito mortality and a decrease in WEEV dissemination. The highest proportion of surviving *Cx. pipiens quinquefasciatus* that became infectious at 20°C occurred 20 days after initial exposure with the virus ($t = 23.5$ days). Under these conditions, the peak of infectiousness of *Cx. tarsalis* with WEEV would be reached by the second blood meal. For WNV in *Cx. pipiens quinquefasciatus*, it would be reached by the fifth blood meal, a point not likely to be reached by mosquitoes in nature.

At 30°C, dissemination progressed faster for both viruses, and *Cx. tarsalis* infectiousness with WEEV peaked by the second blood meal ($t = 6.5$ days). Infectiousness for *Cx. pipiens quinquefasciatus* with WNV increased steadily until all surviving mosquitoes became capable of virus transmission by the fourth blood meal.

Considering both temperature regimens, the peak of the proportion of surviving mosquitoes that were infected occurred during the first blood meal, but the peaks in the proportion of surviving mosquitoes that were infectious occurred days after. Therefore, the peaks of infection and infectiousness do not coincide. Thus, variability in dissemination over time causes the proportion of infectious mosquitoes to not be a constant fraction of the number of infected mosquitoes, even under these simplified conditions.

**Model II: Relationship between population infection prevalence and estimated infection rate.** The process of estimating infection rates using samples taken from a mosquito population was modeled using numerical simulations. The simulations consisted of sampling a hypothetical population of mosquitoes with a known infection rate, randomly grouping the sampled individuals into pools, and simulating viral assays with different abilities to detect virus (Figure 4). Variations were introduced on sample size (200, 2000), pool size (20, 50) and in the ability of the assay to detect the virus (low, high). All the simulations were conducted using R software 2.7.0.

Hypothetical mosquito populations consisted of 100,000 individuals. Each population had two attributes: proportion of infected mosquitoes and a distribution of virus titers among infected mosquitoes. Four proportions of infected mosquitoes in the population were considered: 1/1,000, 5/1,000, 10/1,000, or 15/1,000. Two virus titer distributions were used. Distribution 1 had titers between $10^{1.5}$ and $10^4$ PFU/body, and Distribution 2 had titers between $10^2$ and $10^6$ PFU/body.

---

**Figure 2.** Dissemination of virus in mosquitoes at different temperatures. Each panel shows two dissemination curves with values taken from the literature (references cited in parenthesis in legends), and a third curve of values that were estimated and used as parameters (D) for Model I. (A and B) West Nile virus dissemination in *Culex pipiens* and *Cx. pipiens quinquefasciatus*. (C and D) Western equine encephalomyelitis virus dissemination in *Culex tarsalis*. For the *Cx. tarsalis*–western equine encephalomyelitis virus system, percentages correspond to oral transmission and not dissemination, and the values were approximated from figures in the original publications.
These patterns approximate those observed in *Cx. pipiens* 4–30 days post-infection with WNV at 18–20°C (Distribution 1) and at 26–30°C (Distribution 2). Titers were assigned randomly to each infected mosquito in the population using a uniform distribution. By randomly assigning titers, we were assuming that mosquitoes at any point in the virus extrinsic incubation period could be collected at any particular time. The two distributions overlap, but this has also been observed in *Cx. tarsalis* infected with WEEV and WNV. Our goal was to determine how these titer distributions interact with the virus detection...
ability of the assays to influence the resulting value of the estimated infection rate.

Because simple random sample sizes of 200 and 2,000 were drawn with replacement, each repeated sample was taken from the original population. A total of 1,000 samples of each size were taken from each hypothetical population. It was assumed that only females were sampled and that they were all equally likely to be collected regardless of their age, gonotrophic status or feeding status (empty, blooded). These sample sizes were selected to represent small and large samples that could be common during surveillance. The probability of the sample containing infected mosquitoes increases asymptotically for small and large samples as the proportion of infected mosquitoes in the population increases.

Each sample was divided into groups or pools, as is usually done in surveillance studies to screen large numbers of mosquitoes, and each pool was tested for the presence of virus. Two pool sizes were considered, 20 and 50 mosquitoes. Each sample of 200 mosquitoes was divided into 4 pools of 50 or 10 pools of 20. Each sample of 2,000 mosquitoes was divided into 40 pools of 50 or 100 pools of 20. The mosquitoes were randomly assigned to each pool.

Each simulated pool was individually evaluated for the presence of virus. Plaque-forming units of virus were assumed to be individual discrete units, and if more than one infected mosquito was present in the pool their titers were summed. The total PFUs in the pool was then divided by 2.5 to generate a PFU/mL concentration for each pool to simulate typical homogenization and extraction methods (using 2.5 mL of diluent as for the VecTest).

Two viral assays with different virus detection ability were simulated. The first one was an assay that could only detect virus at concentrations $\geq 10^{4.8}$ PFU/mL. This was called the low detection ability assay. The second one was an assay that detected virus at concentrations $\geq 10^{5.5}$ PFU/mL. This was called the high detection ability assay. The concentration threshold for the low detection ability assay is based on reports that WNV in a pool can be detected up to a minimum concentration of $10^{3.8}$ PFU/mL with the VecTest or $10^{3.7}$ PFU/mL with RT-PCR. The threshold for the high detection ability assay was based on reports that the VecTest and RT-PCR can detect virus in pools of 50 Cx. tarsalis containing only one infected female, which had been inoculated with 100 PFU of virus and held at 28°C for 3 days.

To explore which sample sizes, pool sizes, or virus detection assays were more likely to produce estimated infection rates close to the true population infection rate, their MLE frequency distributions were compared by using descriptive statistics (within population comparisons). Median, minimum and maximum values, and percentage of samples that produced estimates under the population infection rate (percentage of underestimation), were calculated for each frequency distribution. Even MLE values of 0.01 infected mosquitoes/1,000 below the population infection rate were considered underestimates.

To determine if MLE distributions were different between populations, pairwise comparisons of distributions were conducted by using a Kolmogorov-Smirnov test (between population comparisons). The pairwise comparisons tested differences between MLE frequency distributions for populations with 1/1,000 infected mosquitoes versus MLE frequency distributions for populations with 5/1,000 infected mosquitoes. It follows that the other comparisons were between 1/1,000 versus 10/1,000, 1/1,000 versus 15/1,000, 5/1,000 versus 10/1,000, 5/1,000 versus 15/1,000, and 10/1,000 versus 15/1,000. A Bonferroni correction for multiple comparisons was applied and the overall probability to test the null hypothesis of equal distributions was 0.05/6 = 0.0083. Comparisons were made between distributions that were generated with the same sample size, pool size, and virus detection method. This comparison was necessary because those factors determine the number of outcomes in the distribution, and for the Kolmogorov-Smirnov test one needs the same number of outcomes in both groups being compared.

**Results for Model II.** Frequency distributions with a total of 1,000 MLE values were obtained after simulated sampling, pooling, and testing of mosquitoes from hypothetical populations. A total of eight frequency distributions were obtained for each population (Figure 5).

The descriptive statistics for the within population comparisons of the frequency distributions are shown in Tables 2 and 3. When hypothetical populations followed viral titer Distribution 1 ($10^{5.5} \ldots 10^9$ PFU/mosquito body), the median MLE values fell well below the population infection rate that they were attempting to estimate (Figure 6A and B) and underestimation was always $> 80\%$ (Table 2). When simulated tests with low virus detection ability were used with samples from a population with titer Distribution 1, virus detection failed in a large number of positive samples and most MLE had values of zero (Table 2). For example, we knew that 17.4% of the 1,000 samples of size 200 taken for the population with
infection rate 1/1,000 and Distribution 1 contained at least one infected mosquito. However, none of those samples produced positive pools after the simulation with the low detection ability assay and the MLE values were all 0.

When hypothetical populations had a viral titer Distribution 2 ($10^2$–$10^6$ PFU/mosquito body), median estimated infection rates (MLE) were relatively closer to the population infection rate, especially when a test with high viral detection ability was simulated (Table 3 and Figure 6C and D). However, underestimation remained the rule. The frequency distributions with median MLE closer to the true value were obtained when population infection rates were 5/1,000 and the high viral detection ability test was simulated. Sample sizes of 2,000 tended to produce frequency distributions with median MLEs that increased with the population values, especially when combined with high viral detection ability tests (Figure 6D).

These results suggest that viral titer interacts with the ability of the test to detect virus, affecting the distribution of MLE outcomes: lower titers are largely undetected by assays with low ability to detect virus. Also, simulated large sample sizes and assays with high detection ability produced median MLE values that best approximated the population infection rate in the hypothetical populations (Figure 6D). The two pool sizes

---

**Table 2**

Median maximum likelihood estimator values, minimum and maximum values, and percentage of maximum likelihood estimator values under the population infection rate (percentage underestimations) for 32 frequency distributions calculated for hypothetical mosquito populations with viral titer Distribution 1.

<table>
<thead>
<tr>
<th>Population infection rate</th>
<th>Sample size</th>
<th>Pool size</th>
<th>Low virus detection ability</th>
<th>High virus detection ability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>MLE</td>
<td>Minimum–maximum</td>
</tr>
<tr>
<td>1/1,000</td>
<td>200</td>
<td>20</td>
<td>0</td>
<td>0–0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0</td>
<td>0–0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20,000</td>
<td>0</td>
<td>0–0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0</td>
<td>0–0</td>
</tr>
<tr>
<td>5/1,000</td>
<td>200</td>
<td>20</td>
<td>0</td>
<td>0–0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0</td>
<td>0–0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2,000</td>
<td>0</td>
<td>0–0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0</td>
<td>0–0</td>
</tr>
<tr>
<td>10/1,000</td>
<td>200</td>
<td>20</td>
<td>0</td>
<td>0–4.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0</td>
<td>0–4.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2,000</td>
<td>0</td>
<td>0–0.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0</td>
<td>0–0.50</td>
</tr>
<tr>
<td>15/1,000</td>
<td>200</td>
<td>20</td>
<td>0</td>
<td>0–4.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0</td>
<td>0–4.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2,000</td>
<td>0</td>
<td>0–0.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0</td>
<td>0–0.50</td>
</tr>
</tbody>
</table>

*Two of the samples produced all positive pools and were excluded because maximum likelihood estimator (MLE) values cannot be calculated with all positive pools.
produced almost identical MLE frequency distributions. Thus, the effect of pool size in these simulations was minimal.

Kolmogorov-Smirnov pairwise comparisons tested if the MLE frequency distributions were significantly different as the population infection rates increased (between population comparisons). Significant differences in the MLE frequency distributions were detected only between populations with infection prevalence of 1/1,000 and 10/1,000 and between 1/1,000 and 15/1,000, but only when the frequency distributions came from samples sizes of 2,000 and when the populations had titer Distribution 2 (Table 4). Thus, estimated infection rates only reflected changes in the actual population infection when the difference in prevalence was ≥10 times larger and when titers in mosquitoes were high and large samples were taken.

**DISCUSSION**

The results of this study suggest that estimated infection rates are more likely to reflect changes in infection prevalence in the mosquito population only when the sampling and virus testing are conducted in the best available conditions (i.e., large samples paired with assays that can detect virus at low concentrations). However, even under those conditions, infection estimates from samples are highly likely to underestimate the infection rate in the mosquito population. More importantly, these estimated infection rates are not straightforward indicators of changes in the risk of arbovirus transmission. This finding is caused by the variation among mosquitoes to complete the EIP and disseminate virus as a consequence of temperature, the time elapsed after initial infection, and the particular mosquito and virus species. For example, consider the cohort in Model I. If mosquitoes in the cohort were sampled shortly after the initial blood meal and tested under the best available conditions, the estimate of infection prevalence would be approximately 1%, the hypothetical population value. However, the risk of transmission at this point is null because there has not been any virus dissemination. Conversely, if the surviving mosquitoes were sampled around the time of the second or third blood meals, the proportion of infectious mosquitoes would be greater, but there would not be an associated significant increase in the estimated infection rate (still at 1%). Thus, when similar estimated infection rates are obtained from mosquito samples taken at different times or places, these estimates may not carry the same information about risk of virus transmission. In a larger context, if similar values of estimates of infection are obtained for the same mosquito–virus system at different periods, these estimates may not be equivalent because of variation introduced by differences in mosquito survival and environmental temperature over time. Similar infection rates estimated for two mosquito–virus systems would also be not comparable because of variations in the rate of virus dissemination associated with the biology of the system.

Model I considers only a single cohort of mosquitoes, but a population would be composed of numerous cohorts of mosquitoes at different ages. Depending on the species and environmental conditions, recruitment to the population can occur infrequently, in large emergences or steadily over time. Our model would be more appropriate to situations with infrequent, large emergences and illustrates the problem of interpretation of infection rates in the field. We anticipate that continual recruitment to a population, and the resulting variation in timing of blood meals, would increase the discrepancy between the infected and infectious populations. Further research is needed on populations with a dynamic age structure and its effect on infection dynamics.

To have a more direct measure of infectiousness, field-collected mosquitoes could be tested individually in the laboratory for virus dissemination to the salivary glands. Also, traps baited with susceptible hosts could be placed in the field to capture biting mosquitoes, test them for virus, and later test the host for infection. In that way, one would have an estimate of mosquito infection paired with a measure of transmission.18,39 However, these methods are seldom feasible, practical, or affordable for long-term surveillance programs.

Thus, estimated infection rates are necessary, and there is a large body of research dealing with ways to increase the

---

**Table 3**

Median maximum likelihood estimator values, minimum and maximum values, and percentage of maximum likelihood estimator values under the population infection rate (percentage underestimations) for 32 frequency distributions calculated for hypothetical mosquito populations with viral titer Distribution 2

<table>
<thead>
<tr>
<th>Population infection rate</th>
<th>Sample size</th>
<th>Pool size</th>
<th>Low virus detection ability</th>
<th>High virus detection ability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Median MLE</td>
<td>Minimum-Maximum</td>
</tr>
<tr>
<td>1/1,000</td>
<td>200</td>
<td>20</td>
<td>0</td>
<td>0–4.99</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0–4.93</td>
</tr>
<tr>
<td>2,000</td>
<td>200</td>
<td>20</td>
<td>0.50</td>
<td>0–2.03</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0–2.08</td>
</tr>
<tr>
<td>5/1,000</td>
<td>200</td>
<td>20</td>
<td>0</td>
<td>0–23.67</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0–20.20</td>
</tr>
<tr>
<td>10/1,000</td>
<td>200</td>
<td>20</td>
<td>2.03</td>
<td>0–6.34</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>2.08</td>
<td>0–7</td>
<td>97.3</td>
</tr>
<tr>
<td>15/1,000</td>
<td>200</td>
<td>20</td>
<td>4.99</td>
<td>0–23.67</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>4.93</td>
<td>0–20.20</td>
<td>82.6</td>
</tr>
<tr>
<td>2,000</td>
<td>200</td>
<td>20</td>
<td>4.14</td>
<td>0.50–9.23</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>4.39</td>
<td>0.50–9.21</td>
<td>100</td>
</tr>
<tr>
<td>50</td>
<td>4.93</td>
<td>0–20.20*</td>
<td>91.9</td>
<td>11.35</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.90</td>
<td>2.03–14.20</td>
<td>100</td>
</tr>
<tr>
<td>2,000</td>
<td>50</td>
<td>7</td>
<td>1.54–14.51</td>
<td>100</td>
</tr>
</tbody>
</table>

*Some samples (<5%) produced all positive pools and were excluded because maximum likelihood estimator (MLE) values cannot be calculated with all positive pools.
reliability of our estimates of infection. One common recommendation is to take large samples of mosquitoes because infection is a rare event and a large sample increases the probability of collecting an infected mosquito, improving the accuracy of the estimated infection rate. In the laboratory, grouping mosquitoes in small pools is recommended to minimize the random error of the estimated infection rate because small pools approach individual mosquito testing. However, smaller pools would make the testing less cost-effective and may be beyond the reasonable limits of public health infrastructures in many areas. The limits of pool size would not only be determined by budget and facilities but also by the maximum pool size recommended for a particular assay on the basis of its chemistry. Results from our numerical simulations point to the ability of an assay to detect virus, together with the sample size, as important factors affecting the reliability of the estimates of infection. When viral titers in the population were low (Distribution 1) or the population infection rate was low (1/1,000), the viral assay with low virus detection ability (meaning it needed relatively high concentrations of virus to produce a positive result) frequently underestimated (or failed to detect) the population infection prevalence. Therefore, using assays with the highest known detection ability is recommended. Because ideal conditions to estimate infection rates (large samples, small pools, and assays with high ability to detect virus) are difficult to achieve, it is of utmost importance to understand how biases accumulate throughout the estimation process and to be aware of them when making assessments or decisions about transmission risk based on infection rates.

One of the most important goals when monitoring infection rates is to detect changes that could be related to increased risk of arbovirus transmission to animals and humans. To determine if changes are unusual, it is necessary to calculate a baseline infection rate. Increases in infection rates could occur prior to

---

**Figure 6.** Results of Model II showing the median outcomes of the estimated infection rate (maximum likelihood estimator [MLE]) frequency distributions (see Tables 2 and 3) obtained after simulated sampling and testing of mosquitoes coming from hypothetical populations with increasing infection prevalence (1, 5, 10, and 15 infected mosquitoes/1,000). Plots A and B correspond to median MLE outcomes for populations with viral titer Distribution 1. Because the estimation usually did not differ between different pool sizes, the lines are sometimes superimposed. Plots C and D correspond to median MLE outcomes for populations with viral titer Distribution 2. A gray solid line indicating perfect agreement between population prevalence and estimated infection rate is included for comparison purposes.
or during outbreaks due to virus amplification events when biologi-

cal and environmental factors are ideal for large numbers of mosqui-
tes to transmit virus to a large population of sus-
ceptible amplification hosts.\textsuperscript{44} Infec-
tion rates could also change after viremic hosts are intro-
duced into an area.\textsuperscript{45} The reliability of the
methods that we use to estimate infection rates, and the
range of the baseline information that we have, will determine
our capacity to identify unusual changes in virus activity.

However, in addition to baseline estimated infection rate
data, information on other mosquito population parame-
ters and environmental variables is necessary to determine
if unusual arboviral transmission is occurring. Monitoring
changes in the abundance of parous females, changes in the
relative abundance of the total mosquito population, tem-
perature, and rainfall patterns, among others, are necessary
indicators that should be used in conjunction with estimated
infection rates in assessing risk.\textsuperscript{2} These factors should also be
investigated for extended periods to determine baseline pat-
terns during transmission season and then compare these val-
ues to periods of epizootic and epidemic transmission.

The use of multiple criteria for risk assessment is the cur-
cent procedure in some areas, such as California and Florida.
These surveillance programs have a set criteria used to deter-
mine the level of risk of mosquito-borne arbovirus transmis-
sion, and have defined appropriate responses according to
the assessed level of risk.\textsuperscript{24,46} Data on factors such as weather
conditions, vector abundance, mosquito infection rates, sen-
tinel chicken seroconversion, dead bird detection, and human
cases are compared to historical data to determine the level of risk (e.g., normal season versus emergency) and the actions
that need to be taken (public announcements, mosquito control, etc.).\textsuperscript{23,46}

We acknowledge that the risk of transmission is not only a
function of the proportion of infectious mosquitoes in a popu-
lation, but also a function of their abundance and biting pref-
erences. Recently, new measures of risk assessment have been
proposed and could be more useful when comparing the risk
in different areas or seasons, or the risk arising from different
mosquito species. These are 1) the density of infected mosqui-
toes, resulting from the product of mosquito abundance and
the estimated infection prevalence;\textsuperscript{19,67} 2) the probability that a
mosquito species will infect a mammal,\textsuperscript{48} which is calculated as the
product of abundance, proportion of blood meals taken
from mammals, the estimated infection rate, and the fraction
of infected mosquitoes that will subsequently transmit virus by
bite; and 3) the vector index, which considers multiple species
and is calculated as the summation of the product of the aver-
age number of mosquitoes of each species per trap times the

<table>
<thead>
<tr>
<th>Sample size and viral titer</th>
<th>Ability to detect virus</th>
<th>Pool size</th>
<th>1/1,000 vs. 10/1,000</th>
<th>D</th>
<th>P</th>
<th>D</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000, Distribution 2</td>
<td>Low</td>
<td>20</td>
<td>0.46</td>
<td>0.00081</td>
<td>0.51</td>
<td>0.00012</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.33</td>
<td>0.07 (NS)</td>
<td>0.47</td>
<td>0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>20</td>
<td>0.32</td>
<td>0.04 (NS)</td>
<td>0.46</td>
<td>0.00081</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.40</td>
<td>0.02 (NS)</td>
<td>0.43</td>
<td>0.007</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a} Results support significant differences among maximum likelihood estimator frequency distributions among 2000 and 2000 vs. 2000. Distribution 2 Low 20 0.46 0.00081 0.51 0.00012 NS = no significant difference.

In addition to the factors examined here, there are many
other factors that may influence the relationships between
infection rates and risk of transmission. For example, the risk
of transmission of a particular virus can change when multiple
mosquito vector species are considered together. Virus strains
could also vary in their dissemination rates under different
temperatures,\textsuperscript{50} which would add another level of complex-
ty to the relationship. In this study, we did not examine other
environmental factors such as larval rearing conditions that
can affect the susceptibility of mosquitoes to become infected
and/or infectious.\textsuperscript{51} In our second model, we simulated two
viral titer distributions, but empirical data on the distribution
of titers in field-collected mosquitoes is limited. One study
documented coarse variations in titers among different virus
and mosquito species over space and time, and suggested that
monitoring high titers of virus in mosquitoes in the field could
help to identify species with higher vector competence and to
observe the effects of temperature in the progression of the
virus amplification cycle.\textsuperscript{52} In our study, we assumed that viral
titers can be summed when more than one infected mosquito
is present in a pool, but it is necessary to investigate how hav-
ing multiple positive mosquitoes in a pool determines virus
concentration. We did not consider other factors that could
introduce biases and error in estimation, such as trapping
biases and the false-positive detection of virus.

In Florida, field experiments using sentinel chickens to esti-
mate the rate of virus transmission by \textit{Cx. nigripalpus} showed
that transmission is low in relation to estimated infection
rates,\textsuperscript{38,39} and many factors that we studied here were consid-
ered to explain this occurrence. These field studies support
the idea that incorporation of information on the mosquito
population (age structure, abundance and estimated infection
rate), virus (strain, species-specific dissemination rates), and
the environment (temperature) are necessary to make assess-
ments of the risk of virus transmission.

The results from our simplified models remind us that esti-
mates of mosquito infection prevalence may not be directly
related to the risk of transmission of arboviruses to animals
and humans. The general assumption that increases in infection
rates indicates increases in the risk of arbovirus transmission
might not be reliable as supported by field data. It is impor-
tant to remember that 1) there are biases that sampling and
virus testing methods introduce in our estimates, which will
cause them to generally underestimate the population infec-
tion rate; 2) similar estimated infection rates across locations,
virus (strain, species-specific dissemination rates), and
the environment (temperature) are necessary to make assess-
ments of the risk of virus transmission.

The results from our simplified models remind us that esti-
mates of mosquito infection prevalence may not be directly
related to the risk of transmission of arboviruses to animals
and humans. The general assumption that increases in infection
rates indicates increases in the risk of arbovirus transmission
might not be reliable as supported by field data. It is impor-
tant to remember that 1) there are biases that sampling and
virus testing methods introduce in our estimates, which will
cause them to generally underestimate the population infec-
tion rate; 2) similar estimated infection rates across locations,
time, and for different mosquito and virus species might not
indicate similar risk of arbovirus transmission because of vari-
ations in dissemination rates; 3) arbovirus baseline data are
necessary to define unusual changes in transmission activity;
and 4) use of other surveillance indicators in conjunction with
estimated infection rates is necessary to fully assess the risk of
arboviral transmission.
Received June 12, 2009. Accepted for publication February 8, 2010.

Acknowledgments: We thank K. Pesko for collaborating during the initial work in this study; Drs. M. Reiskind, J. F. Day, and L. P. Lounibos for reviewing the manuscript and providing many useful suggestions; and Carol Thomas for assisting in manuscript preparation.

Financial support: This study was supported by the National Institutes of Health (grant AI42164).

Authors’ addresses: Dulce M. Bustamante, Escuela de Biología, Universidad de San Carlos, Ciudad Universitaria, Guatemala City, Guatemala. E-mail: dulce_mariab@hotmail.com. Cynthia C. Lord, Florida Medical Entomology Laboratory, Department of Entomology and Nematology, University of Florida, Vero Beach, FL. E-mail: clord@ufl.edu.

REFERENCES


