

## **Monitoring of prevalence of infection in vectors in an era of elimination**

Thomas R. Unnasch, Ph.D

Charles R. Katholi, PhD

Sampling of vectors for the presence of filarial parasites is desirable because it provides a current snapshot of the intensity of parasite - vector contact and transmission. This is in contrast to methods relying upon patent infections in the host population which provide a delayed readout of transmission intensity. However, in areas subject to a successful control program, the prevalence of infected (or infective) vector insects is very low. This presents several technical obstacles to obtaining an accurate estimate of the prevalence of infection in the vector population. First, when the prevalence of infection in the vector population is low, the specificity of the assay used to detect the presence of the parasite must be extremely high in order to ensure an acceptable positive predictive value. Second, a large number of insects must be screened to obtain an accurate estimate of infection, which in areas under control is a very rare event. Highly sensitive and specific molecular assays, coupled with pool screening methods can overcome these technical obstacles.

How samples are collected for pool screening is an essential component in the process of monitoring infection in a vector population. Sampling methods can be targeted to those areas that are most likely to contain infected insects. Such a targeted sampling scheme will provide the most sensitive indicator of ongoing transmission. However, data collected using a targeted sampling method will over-estimate the actual transmission intensity. To obtain an accurate estimate of transmission intensity, more global sampling methods are required. There are a number of sampling methods which can be used when attempting to measure the prevalence of an infection over a large geographic area. When the infection is known to be restricted to certain well identified sub-regions then either systematic sampling or cluster sampling is often employed. Since the sampling method affects the analysis of the data, it is important to consider the sampling design and the subsequent analysis issues together prior to gathering data. Part of any sampling design is deciding whether to use *Simple Random Sampling* or some kind of *Probability Sampling* (e.g. Probability proportional to size). In simple random sampling, the probability that any unit will be chosen to be sampled is equal to any other, and no special estimation procedures are required. Current algorithms that calculate prevalence of infection in the vector population based upon pool screening (e.g. PoolScreen v2.0) are based upon the assumption that simple random sampling has been used to collect the samples to analyze. On the other hand, if probability sampling is chosen, then special estimators are required which reflect the selection probabilities of each pool processed. In addition, software is required to draw the sample for collection and analysis. The software is necessary, since after each cluster is selected then the selection probabilities for the remainder change and must be recalculated. Furthermore, obtaining point estimates for the prevalence of infection from probability samples sampling protocols is computationally difficult, and methods for deriving variance estimates and confidence intervals from pool screened data derived from samples collected by probability based methods have not been developed. Thus, simple random sampling methods are highly preferred over probability based methods in situations where the samples collected are to be screened using pool screening.