**ONLINE SUPPLEMENT: Estimating the percent of true infections among source patients**

**METHODS**

To estimate the true frequency of infection among all specimens submitted from source patients to the seven participating laboratories, it is necessary to correct the frequency of positive results as reported by laboratories (observed percent positive) for the likely occurrence of false positive and false negative test results.  Generally speaking, the observed percent positive reflects the frequency of true infections modified by the sensitivity and specificity of the assay used, as indicated by the following equation:

1. Observed % Positive =[% True Infection x Assay Sensitivity] + [(1-% True Infection) x (1-Assay Specificity)]

Given the observed percent positive and the sensitivity and specificity of the assay, it is possible to back-calculate the percent of true infections among the samples tested. For Lyme disease, however, the sensitivity of serologic testing varies considerably depending upon whether the patient has localized or disseminated infection.  Therefore, it is necessary to modify the equation as follows to account for different stages of illness among source patients:

Observed % Positive = [% True Infection x  Sensitivitylocalized  x % Localized Disease] + [% True Infection x  Sensitivity disseminated x % Disseminated Disease)] + [(1-% True Infection) x (1-Specificity)]

Substituting the equivalent value (1-% Localized Disease) for % Disseminated Disease, this equation can be simplified to:

Observed % Positive = [% True Infection x Sensitivitylocalized  x % Localized Disease] + [% True Infection x Sensitivity disseminated x (1-% Localized Disease)] + [(1-% True Infection) x (1-Specificity)]

Values for Sensitivitylocalized, Sensitivitydisseminated and Specificity for equation (c) can be derived from the published literature and the observed percent positive can be determined empirically from the laboratory survey results.  The remaining two unknowns, percent true infection and percent localized disease, cannot be determined from a single equation because there are many possible combinations of the unknowns that would yield the same observed percent positive.  It is possible, however, to derive two independent equations, one relating these unknowns to the sensitivity, specificity, and observed percent positive for first tier (EIA) testing, and a second for two-tiered testing.  As with an algebraic problem involving two equations and two unknowns, there are few combinations of the unknowns that can match the observed results for both assays simultaneously.

To identify these values, we iteratively calculated for each equation a predicted percent positive for all combinations of percent true infection and percent localized disease based on the sensitivity and specificity of the test.  This was done for combinations of percent true infection ranging from 6-20%, and for percent localized disease ranging from 0-100%.  These values of predicted percent positive were then compared with the observed percent positive for each assay, based on reported laboratory survey results for the four endemic states.  Differences between predicted and observed results were summed for both assays to determine the combination of percent true infection and percent localized disease that best predicted the observed data for both assays simultaneously.  We did not include an equation for Western blot standalone data in this analysis, as a portion of the source population for these assays have likely had pre-testing by ELISAs, which would increase the percent true infection in this specific group.

To estimate the total number of infections nationwide among source patients, we applied the percent true infection (derived from the above process) to the total number of specimens tested by participating laboratories. This calculation assumes that the four states from which the true rate was derived are representative and that only one specimen is submitted for each patient.

**RESULTS**

The frequency of positive two-tiered and EIA/ELISA tests reported by five national laboratories for specimens from CT, MD, MN, and NY are presented in Table 3. Also included are the parameters used to estimate the percent of true infections [[18-23](#_ENREF_18)]. The difference between observed and predicted percent positive results, based upon different combinations of percent true infection and percent localized disease, are presented in Figure 1. An estimated percent true infection of 12%, with 86% of samples coming from patients with localized disease, yields the closest match to observed results for both assays combined. At these values, the difference between the observed and predicted percent positive values for the two-tiered test is 0.03% and for EIA/ELISA tests is 0.00%. Also shown in Figure 1, a reasonable fit with the observed data could only be attained with a percent true infection of between 11 and 13%, even when the percent early disease ranged from 66 to 100%. All other combinations resulted in substantially greater differences between the observed and predicted percent positive values. This sensitivity analysis suggests that the estimated percent true infection is robust with respect to the proportion of specimens derived from patients with early versus later stages of infection.

Figure 1. Estimating the percent true infection among samples submitted for Lyme disease testing using the total difference between observed and predicted test results in four endemic states (CT, MD, MN, NY).