A COMPARATIVE ANALYSIS OF LIKELIHOOD RATIO STATISTICAL RANKING TO EXPECTED MATCH RATIOS AND ESTIMATED KINSHIP RATIOS FOR FAMILIAL DNA SEARCHING APPLICATIONS

Bonnie Mountain[1], Dr. Stephanie A. Santorico[2] and Dr. Greggory LaBerge[1,3]

[1]Denver Police Crime Laboratory
[2]University of Colorado Denver, Department of Mathematical & Statistical Sciences
[3]University of Colorado Anschutz Medical Center, Human Medical Genetics Program

The Denver Police Crime Laboratory has developed a familial searching system that considers genotypes over all CODIS markers with statistical ranking based on a likelihood ratio (DLR). In this research the DLR was compared to the Scientific Working Group on DNA Analysis Methods (SWGDAM) recommendation of using Estimated Match Ratios (EMR) and Estimated Kinship Ratios (EKR) to assess relatedness. The EMR represents the likelihood of detecting a relative in a DNA database of offenders as compared to an unrelated person, while the EKR is representing the likelihood that two DNA profiles being compared are from related versus unrelated individuals. Similarly, the DLR compares the likelihoods of the two DNA profiles given they are derived from related versus unrelated individuals, however; the SWGDAM recommended statistics and DLR calculations have a few key differences. The SWGDAM recommendations include a normalization factor that is introduced by dividing the EMR and EKR values by the database size whereas the DLR does not perform any additional modifications to the likelihood ratios based on the size of the database being searched or the number of combinations considered as part of a familial search. Low stringency matches are excluded from the EMR/EKR calculations, while all loci being compared are considered in the DLR method. Both methods calculate these likelihood ratios using Caucasian, African American, Southwestern Hispanic, and Southeastern Hispanic allele frequencies but have varying cut-off thresholds. A match under the SWGDAM method is considered useful if the EMR or EKR value from at least one population is greater than 1 and all others are greater than or equal to 0.1. The DLR method employs a less stringent method to determine which match results should be considered suggesting additional follow up on matches with DLR values greater than 100,000 but it does not restrict follow up on matches with lower DLR values.

Several simulated familial searches were performed in our work to determine the overall performance of the SWGDAM and DLR methods. Two simulations focused on the parent-child relationship and two examined sibling pairs. Confusion matrices were used to determine each method's overall performance with a focus on the true positive rate and false positive rate. Preliminary results indicate that the DLR method identified more simulated matches (true positives) compared to the SWGADAM method; however more false positives were also identified in the search results. Conversely, the SWGDAM method excluded more false positives with a negative impact on identification of true positives. Additional simulations were completed to assess the impact of the low stringency exclusion and normalization by database size on the true positive and false positive rates. This poster will discuss the results of these simulation studies and the mathematical analyses completed to support the results.