
Procedure for Statistical Calculations

1.0 Purpose – This document provides the procedure for calculating statistical data on single source, predominant profiles and mixed DNA samples.

2.0 Scope – This procedure applies to Casework analysts and trainees in the Forensic Biology Section who are qualified to perform statistical analysis on DNA profiles and mixtures.

3.0 Definitions

- **Combined Probability of Exclusion (CPE):** The probability that a randomly chosen, unrelated person from a given population would be excluded as a potential contributor to the observed DNA mixture.
- **Combined Probability of Inclusion (CPI):** The probability that a randomly chosen, unrelated person from a given population would be included as a potential contributor to the observed DNA mixture.
- **Intimate Sample:** A biological sample from an evidence item that is obtained directly from an individual's body; it is not unexpected to detect that individual's allele(s) in the DNA typing results.
- **Minimum allele frequency:** A frequency which shall be used for any allele which is observed 5 or less times in the population frequency database (to include variant and off-ladder alleles). The formula for this frequency is $5/2N$ ($2N$ = # of alleles in the population database at that locus).
- **Heterozygote:** Two different alleles at a particular gene locus on homologous chromosomes.
- **Homozygote:** Same alleles at a particular gene locus on homologous chromosomes.
- **Product Rule:** A calculation that can be used if a locus is inherited independently of the other loci. The calculation can be made by multiplying each individual genotype frequency together.
- **Random Match Probability (RMP):** The chance that a randomly selected unrelated individual from a population will have an identical STR profile to that obtained from a forensic sample.

4.0 Equipment, Materials and Reagents

- CODIS software program (Analyst Workbench) located on the CODIS computer with Print to PDF software.

5.0 Procedure

5.1 Random Match Probability (RMP)

5.1.1 RMP shall be calculated on all single source or single predominant profiles with the exception of intimate samples where the question sample matches/is consistent with the owner or a consensual partner.

5.1.2 In order to use RMP, the Forensic Scientist must have a minimum of 3 loci for which stats can be provided. Note: Population frequency data for D2S1338 and D19S433 have not been determined for the NC populations and are not used for calculations. Amelogenin frequencies are not used for calculations.

5.1.3 The formulas used in the calculation of the frequency of a DNA profile shall be in accordance with those published in the NRC II guidelines and in PopStats.

5.1.3.1 Heterozygote frequencies: $2pq$ (p and q are the frequencies of the alleles observed).

5.1.3.2 Homozygote frequencies: $p^2 + p(1-p)\theta$, where $\theta = 0.01$ (p is the frequency of the allele observed and θ is a correction factor for expected levels of population substructure in a given population).

5.1.3.2.1 If a locus contains a homozygote allele which falls below the stochastic threshold, that locus shall not be used.

5.1.3.3 Multi-locus frequencies: the product rule shall be used.

5.1.4 Procedure for performing RMP

5.1.4.1 Log onto a CODIS computer.

5.1.4.2 Open the Analyst Workbench program.

5.1.4.3 Click PopStats located in the bottom left.

5.1.4.4 In the Specimen ID box type in the full case number (note: the item number may be entered along with the case number or separately within the comments box).

5.1.4.5 In the Specimen Details block type the alleles.

5.1.4.6 Click Calculate.

5.1.4.7 Print the Forensic/Broward report using the Print to PDF and add it to the FA case record object repository.

5.1.5 See the Forensic Biology Section Procedure for Report Writing for reporting guidelines.

5.2 Combined Probability of Exclusion/Inclusion for DNA mixtures (CPI calculations)

5.2.1 Since CPI is derived from CPE, calculations will be performed for both CPE and CPI. However, only CPI shall be reported.

5.2.2 If a known reference sample cannot be excluded from a mixture profile or from a multiple major profile then CPI shall be calculated (except for intimate samples where the owner or a consensual partner cannot be excluded).

5.2.3 CPI calculations shall be performed only once on a mixture profile except in the instance of multiple major profiles. In that instance, CPI may be performed on the multiple major profile and again on the mixture as a whole for a minor contributor.

5.2.4 If a locus has alleles that fall below the stochastic threshold, that locus shall not be used for statistical weight; however, if a multiple major profile can be determined, the alleles associated with the multiple major profile can be used for calculations.

5.2.4.1 Example: A locus has alleles P, Q, R and S with alleles P, Q and R being above the stochastic threshold and S being below the stochastic threshold. Alleles P, Q and R are interpreted as being from a distinct multiple major profile. CPI calculations can be performed using only alleles P, Q and R.

5.2.5 Formulas: If an evidence sample includes three alleles at a locus (A_1 , A_2 , A_3), then:

5.2.5.1 $P = A_1 + A_2 + A_3$ and

5.2.5.2 $Q = 1 - P$

5.2.5.3 $PE = Q^2 + Q(1-Q)\theta + [2Q(1-Q)(1-\theta)] \quad \theta = 0.01$

5.2.5.4 The probability of exclusion (PE) is calculated at each locus. Then the PEs from the multiple loci are combined through the following equation:

$$CPE = 1 - [(1-PE_1) \times (1-PE_2) \times \dots (1-PE_n)]$$

$$CPI = 1 - CPE$$

5.2.6 Procedure for performing CPE/CPI

5.2.6.1 Log onto a CODIS computer.

5.2.6.2 Open Analyst Workbench program.

5.2.6.3 Click PopStats located in the bottom left.

5.2.6.4 In the explorer window at the upper left, choose Forensic Mixture.

5.2.6.5 In the Specimen ID box type in the full case number (the item number may be entered along with the case number or separately within the comments box).

5.2.6.6 In the Mixture block type in the alleles.

Note: Population frequency data for D2S1338 and D19S433 have not been determined for the NC populations and are not used for calculations. Amelogenin frequencies are not used for calculations.

5.2.6.7 Click Calculate.

5.2.6.8 Print the Probability of Inclusion (Short) report using the Print to PDF and add it to the FA case record object repository.

5.2.7 See the Forensic Biology Section Procedure for Report Writing for reporting guidelines.

6.0 Limitations – N/A

7.0 Safety – N/A

8.0 References

Forensic Biology Section Procedure for Casework DNA Interpretation

Forensic Biology Section Procedure for CODIS

Forensic Biology Section Procedure for Report Writing

National Research Council Committee on DNA Forensic Science (1996) An Update: The Evaluation of Forensic DNA Evidence. Washington, D.C., National Academy Press.

9.0 Records – N/A

10.0 Attachments – N/A

Revision History		
Effective Date	Version Number	Reason
01/03/2013	1	Original Document
03/08/2013	2	Definitions – clarified CPI and CPE definitions; 5.1.3.2 – clarified definition of theta; 5.2.5.2 – changed $Q = 1 - p$ to $Q = 1 - P$