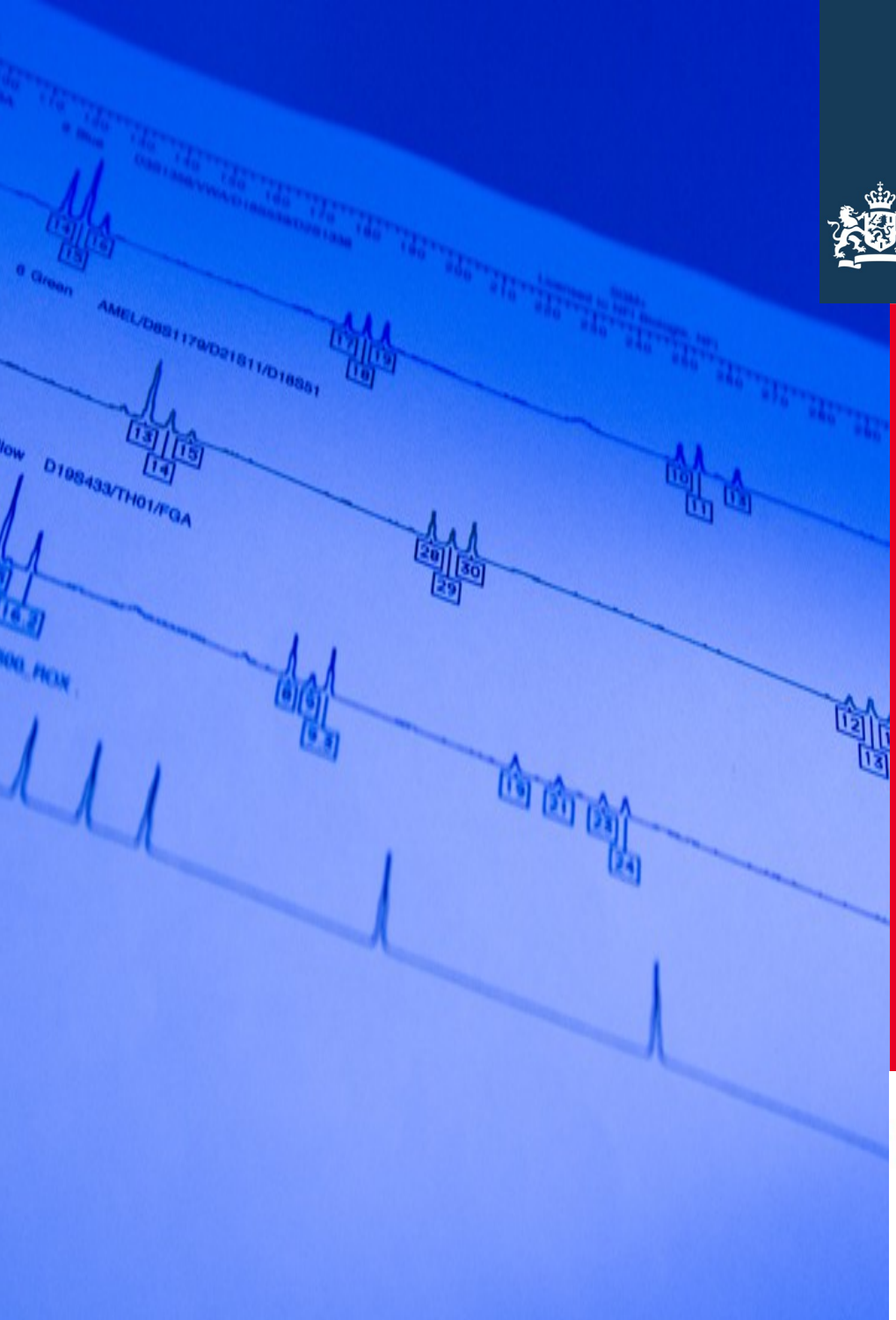

Netherlands Forensic Institute
Ministry of Security and Justice

Statistical approaches for the interpretation of DNA evidence

Hinda Haned
h.haned@nfi.minvenj.nl

Forensic Genomics
Consortium Netherlands

Weight of the evidence

The result of a comparative analysis between a trace and a reference DNA profile must be accompanied by an assessment of the strength of the evidence

Statistical approaches for the weight of DNA evidence

Two categories:

- the 'allele-centric' approach
  - Only the alleles of the mixtures are considered, no genotypes involved
- the 'genotype-centric' approach
  - Genotypes are inferred

Statistical approaches for the weight of DNA evidence

Allele-centric

Probability of exclusion/inclusion (Random Man Not Excluded)

Genotype-centric

- Random Match probability (RMP)
- Likelihood ratios

Probability of exclusion: RMNE

- Mixture: alleles: A, B
- Allele frequencies: $p_A = 0.01$, $p_B = 0.02$

What percentage of the population could be a possible contributor to the observed profile?

- Probability of exclusion, or
- Probability of inclusion

Probability of exclusion: RMNE

What percentage of the population could be a possible contributor to the observed profile?

$$RMNE = (p_A + p_B)^2 = 9 \times 10^{-4}$$
$$PE = 1 - RMNE = 0.9991$$

- The probability of including a person, taken at random from the population is 0.0009
- The probability of excluding a random person is 0.99
- 1 person in 1111 in the population (unrelated to the suspect) could be included

Probability of exclusion: RMNE

What percentage of the population could be a possible contributor to the observed profile?

$$RMNE = (p_A + p_B)^2 = 9 \times 10^{-4}$$
$$PE = 1 - RMNE = 0.9991$$

- The probability of including a person, taken at random from the population is 0.0009
- The probability of excluding a random person is 0.99
- 1 person in 1111 in the population (unrelated to the suspect) could be included

Probability of exclusion: RMNE

Assumptions

- All alleles are visible
- Suspect's alleles are in the sample

Advantages

- Simple to compute and implement
- Easy to explain (jury, judges, lawyers)
- No assumption about the number of contributors

Disadvantages

- Does not make use of all available information
- Cannot be used in LTDNA

RMP: Random Match Probability

RMP

- Assuming a specific number of contributors, what percentage of the population could be a possible contributor to the profile?
- Probability that a random individual will have the same profile than the person of interest, by chance

Random Match Probability

- Sample: alleles A, B
- Suspect: A/B
- Allele frequencies: $p_A = 0.01$, $p_B = 0.02$

- $RMP = Pr(AB) = 2p_A p_B$; $RMP = 4 \times 10^{-4}$
- 1 person in 2500 could be included

Random Match Probability

Advantages

- Simple to compute and implement
- Easy to explain (jury, judges, lawyers)

Disadvantaged

- Assumption about the number of contributors
- Cannot be used in LTDNA

Bayesian probabilities

Above all, it is a way of thinking that established an interpretation framework that follows three rules (Evett & Weir, 1998):

1. To evaluate the uncertainty of any given proposition, it is necessary to consider at least one alternative proposition
2. Scientific interpretation is based on questions of the kind 'What is the probability of the evidence given the proposition?'
3. Scientific interpretation is conditioned not only by the competing propositions, but also by the framework of circumstances within which they are to be evaluated

Likelihood Ratios

Probability of observing the crime sample profile, conditional on two alternative hypotheses explaining the origin of the trace

- Two alternative hypotheses
  - $H_{prosecution}$: Suspect + one unknown are the donors to the sample
  - $H_{defence}$: two unknowns are the donors

Likelihood Ratios

$$LR = \frac{Probability(evidence|H_{prosecution})}{Probability(evidence|H_{defence})}$$

Likelihood Ratios

$$Pr(evidence|H_{prosecution}) = Pr(A, B|Suspect = A, B) = 1$$

$$Pr(evidence|H_{defence}) = Pr(evidence|unknown) = 2p_A p_B$$

$$LR = \frac{1}{2p_A p_B} = 2500$$

The probability of observing the evidence is 2500 times more likely if hypothesis $H_{prosecution}$ is true, compared to its alternative $h_{defence}$

Likelihood Ratios

$$Pr(evidence|H_{prosecution}) = Pr(A, B|Suspect = A, B) = 1$$

$$Pr(evidence|H_{defence}) = Pr(evidence|unknown) = 2p_A p_B$$

$$LR = \frac{1}{2p_A p_B} = 2500$$

The probability of observing the evidence is 2500 times more likely if hypothesis $H_{prosecution}$ is true, compared to its alternative $h_{defence}$

Likelihood Ratios

$$Pr(evidence|H_{prosecution}) = Pr(A, B|Suspect = A, B) = 1$$

$$Pr(evidence|H_{defence}) = Pr(evidence|unknown) = 2p_A p_B$$

$$LR = \frac{1}{2p_A p_B} = 2500$$

The probability of observing the evidence is 2500 times more likely if hypothesis $H_{prosecution}$ is true, compared to its alternative $h_{defence}$

Likelihood Ratios

$$Pr(evidence|H_{prosecution}) = Pr(A, B|Suspect = A, B) = 1$$

$$Pr(evidence|H_{defence}) = Pr(evidence|unknown) = 2p_A p_B$$

$$LR = \frac{1}{2p_A p_B} = 2500$$

The probability of observing the evidence is 2500 times more likely if hypothesis $H_{prosecution}$ is true, compared to its alternative $h_{defence}$

Why should you use LRS?

Advantages

- Hypotheses specification is flexible
- Makes use of more information than any other approach
- Better estimation of the strength of evidence than the RMNE
- Uncertainties on the composition of the DNA sample can be modelled

Disadvantages

- Cannot be computed by hand (need for a software)
- Less intuitive (?) than other methods, more difficult to explain

"L'approche bayesienne oblige à rejeter les visions dichotomiques du monde en juste/faux, possible/impossible. Elle oblige à voir que tout affaire est affaire de nuances de vraisemblance."

"L'approche bayesienne n'est pas simplement une façon de voir les choses (...) c'est en fait la seule démarche correcte pour exprimer de façon explicite et complète la signification d'un résultat de police scientifique. Au premier abord, il faut bien sûr un certain effort pour s'imprégner de sa logique. Mais après quelques exercices, la rigueur, l'élégance et l'efficacité du cadre conceptuel qu'elle offre ne peuvent qu'enthousiasmer ceux qui font l'effort de s'y initier."
(Coquoz & Taroni, 2006)

The Bayesian approach leads to rejecting dichotomous worldviews in right/wrong, possible/ impossible. It leads to considering that everything is a matter of likelihoods.

The Bayesian approach is not just a way of seeing things (...) it is actually the only correct approach to express explicitly and fully the significance of the results of forensic testing. At first glance, it requires some effort to soak up its logic. But after some practice, the rigor, elegance and efficiency of the framework it provides can only inspire those who make the effort of learning it

(attempted translation from Coquoz & Taroni, 2006)

International consensus

ISFG

Available online at www.sciencedirect.com

SCIENCE DIRECT

ELSEVIER

Forensic Science International 160 (2006) 90–101

Forensic Science International

www.elsevier.com/locate/forsciint

DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures

P. Gill [a,*], C.H. Brenner [b], J.S. Buckleton [c], A. Carracedo [d], M. Krawczak [e], W.R. Mayr [f], N. Morling [g], M. Prinz [h], P.M. Schneider [i], B.S. Weir [j]

Recommendation: The likelihood ratio is the preferred approach to mixture interpretation... If the DNA crime stain profile is low level... and or/if the drop-out is possible, then the RMNE method may not be conservative.

International consensus

ISFG

Forensic Science International: Genetics xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

ELSEVIER

DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods

P. Gill [a,b,*], L. Gusmão [c], H. Haned [d], W.R. Mayr [e], N. Morling [f], W. Parson [g], L. Prieto [h], M. Prinz [i], H. Schneider [j], P.M. Schneider [k], B.S. Weir [l]

Recommendation: Probabilistic approaches and likelihood ratio principles are superior to classical methods.

International consensus

- Preferred approach, it does not mean other methods are not valid
- Move towards a more flexible/powerful tool
- Homogenization of interpretation (inter- and intra-laboratories)

Why now?

- LR-based methods are complex
- Need for a software

- Recently, new software have been made available
  - LikeLTD (David Balding, UCL, UK)
  - TrueAllele (Mark Perlin, USA)
  - LRmix (Hinda Haned, Lyon - Den Haag - Oslo)

New interpretation models?

- Models are not new
- Statistical methods existed for decades

NFI: Netherlands Forensic Institute

- LR-based model for he interpretation of DNA profiles
  - Simple profiles
  - DNA Mixtures
  - High template DNA: HTDNA
  - Low template DNA: LTDNA

NFI: Netherlands Forensic Institute

Project Forensic Genetics Consortium Netherlands (FGCN):

- Theoretical developments
- Implementation (software)
- Application on casework

In what follows, we will see how to calculate the likelihood ratios on simple profiles and on mixtures, and when whether there is uncertainty or not on the actual composition of the DNA sample