Validating Software for Probabilistic Genotyping

Simone Gittelson, John Buckleton, Duncan Taylor, and Ian Evett

Disclaimer

Points of view in this presentation are mine and do not necessarily represent the official position or policies of the National Institute of Standards and Technology or the U.S. Department of Justice.

Validating Software for Probabilistic Genotyping

- Tests where H_p is true:

Does the model correctly include a known contributor?

- Tests where H_d is true:

Does the model correctly exclude a known non-contributor?

Forensic Science International: Genetics 16 (2015) 165–171

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Forensic population genetics – original research

Testing likelihood ratios produced from complex DNA profiles

Duncan Taylor[a,b,*], John Buckleton[c], Ian Evett[d]

[a] Forensic Science South Australia, 21 Divett Place, Adelaide, SA 5000, Australia
[b] School of Biological Sciences, Flinders University, GPO Box 2100, Adelaide, SA 5001, Australia
[c] ESR, Private Bag 92021, Auckland 1142, New Zealand
[d] Principal Forensic Services Ltd., London, UK

ARTICLE INFO

Article history:
Received 24 November 2014
Received in revised form 13 January 2015
Accepted 16 January 2015

Keywords:
DNA profile interpretation
Mixtures
Likelihood ratios
Performance tests

ABSTRACT

The performance of any model used to analyse DNA profile evidence should be tested using simulation, large scale validation studies based on ground-truth cases, or alignment with trends predicted by theory. We investigate a number of diagnostics to assess the performance of the model using H_d true tests. Of particular focus in this work is the proportion of comparisons to non-contributors that yield a likelihood ratio (LR) higher than or equal to the likelihood ratio of a known contributor (LR_{POI}), designated as p, and the average LR for H_d true tests. Theory predicts that p should always be less than or equal to $1/LR_{POI}$ and hence the observation of this in any particular case is of limited use. A better diagnostic is the average LR for H_d true which should be near to 1. We test the performance of a continuous interpretation model on nine DNA profiles of varying quality and complexity and verify the theoretical expectations.

1. Introduction

Methods for evaluating DNA profiles have benefitted from

frequentist methods and will call this fraction p but not interpret it as a p-value, which would be more familiar if we were testing some hypothesis. This value of p is relative to a simulation and a model

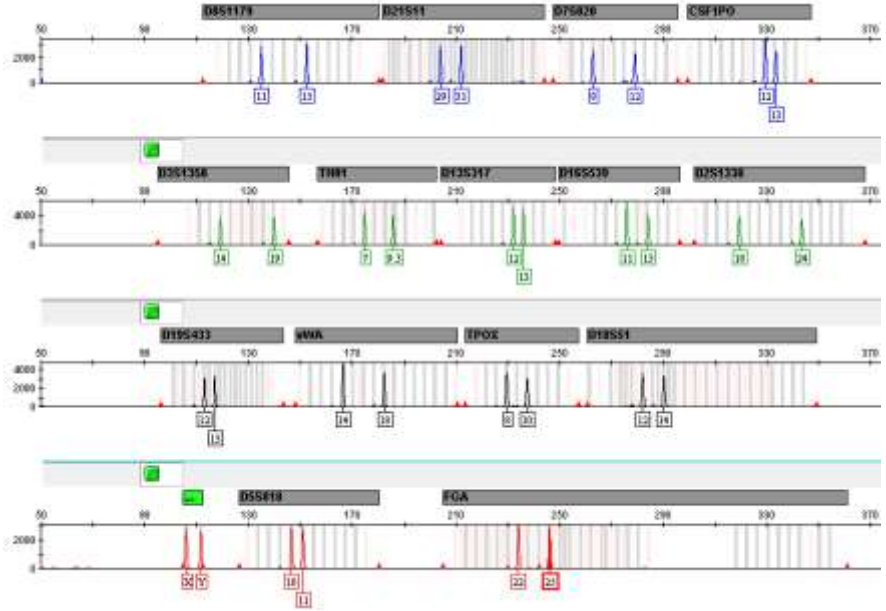
Three Main Points

1) The average value of the likelihood ratios (LRs) obtained for tests where H_d is true is 1.

2) Information on the chance of adventitious matching is provided by the LR.

3) This theory provides a framework for assessing the performance of interpretation models.

Scenario



crime stain

person of interest (POI)

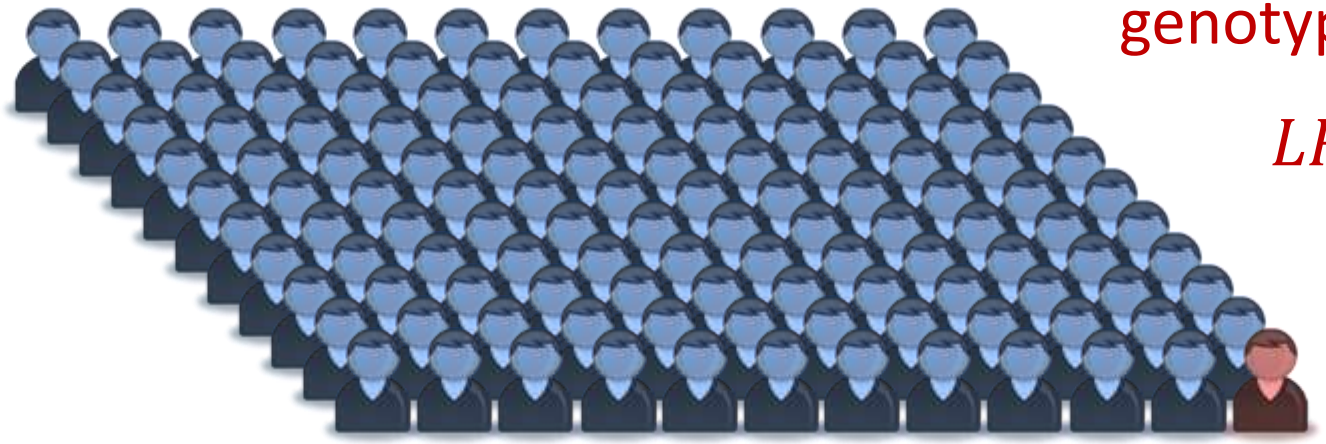
H_p: The POI is the donor.

H_d: An unknown person is the donor.

likelihood ratio $LR = \dfrac{1}{\dfrac{1}{1\ billion}} = 1\ billion$

Our Expectations if H_d is true

1 billion randomly chosen individuals



1 randomly chosen individual has the same genotype as the POI:

$$LR = 1\ billion$$

999,999,999 randomly chosen individuals have a different genotype than the POI:

$$LR = 0$$

Our Expectations if H_d is true

Average value of the LRs:

$$\underbrace{\frac{999{,}999{,}999}{1\ billion} \times 0}_{\text{999,999,999 individuals: } LR = 0} + \underbrace{\frac{1}{1\ billion} \times 1\ billion}_{\text{1 individual: } LR = 1\ billion} = 1$$

1) The average value of the likelihood ratios (LRs) obtained for tests where H_d is true is 1.

$$\sum_i Pr(E_i|H_d) \times LR_i$$

$$= \sum_i Pr(E_i|H_d) \frac{Pr(E_i|H_p)}{Pr(E_i|H_d)}$$

$$= \sum_i Pr(E_i|H_p)$$

$$= Pr(E_1 \cup E_2 \cup \cdots \cup E_n|H_p)$$

$$= 1$$

the expected factor for a wrong hypothesis in virtue of any experiment is 1

Alan Turing

Good I.J. Probability and the Weighing of Evidence. Charles Griffin & Company Limited, London, 1950.

Three Main Points

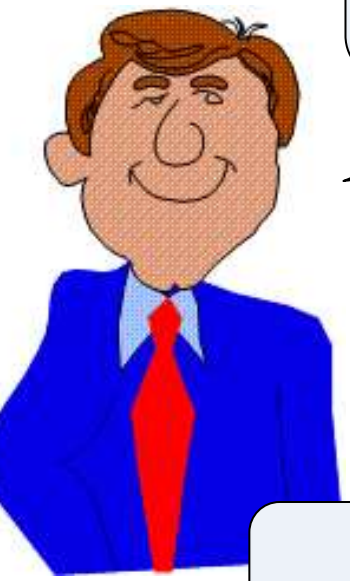
1) The average value of the likelihood ratios (LRs) obtained for tests where H_d is true is 1.

2) Information on the chance of adventitious matching is provided by the LR.

3) This theory provides a framework for assessing the performance of interpretation models.

In court

What were the results of the comparison of the crime stain's DNA with the POI's DNA?



The DNA typing results of the crime stain are $1\ billion$ times more probable to have been obtained if the POI is the donor of this DNA than if an unknown person is the donor of this DNA.

How many other people would adventitiously match this profile?

?

$LR_{POI} = 1\ billion$

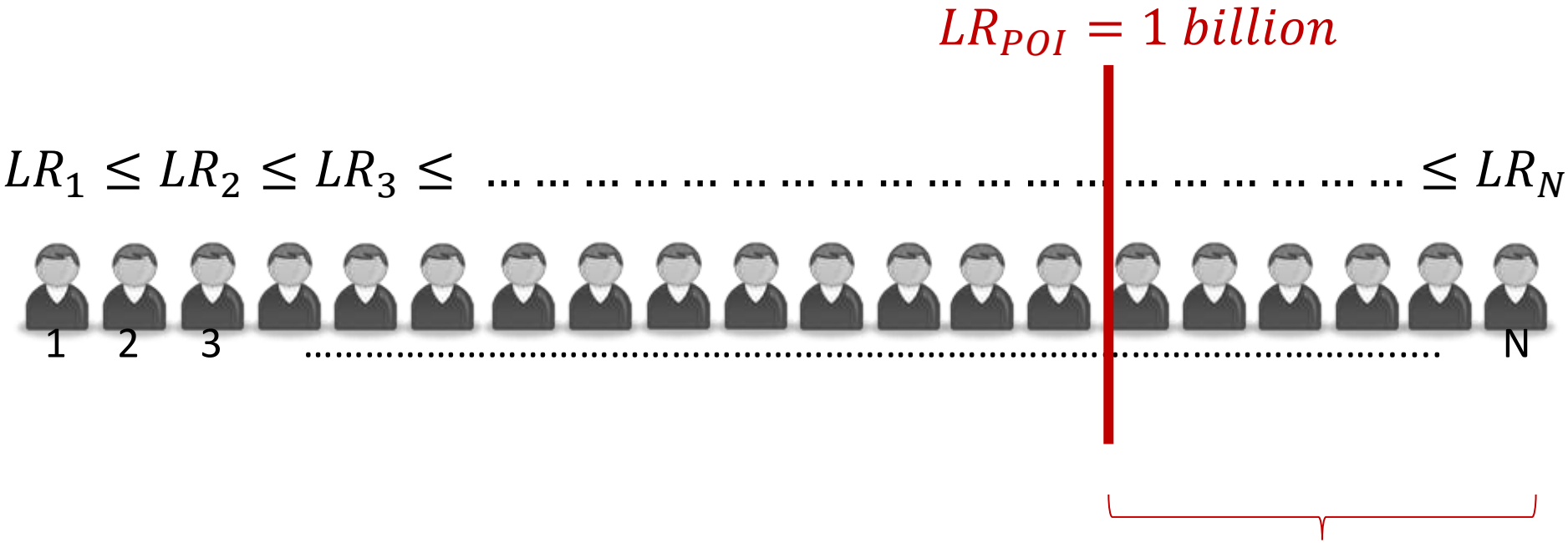
$LR_1 \leq LR_2 \leq LR_3 \leq \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \leq LR_N$

average of LRs smaller than LR_{POI}

average of LRs greater than or equal to LR_{POI}

$= \delta$

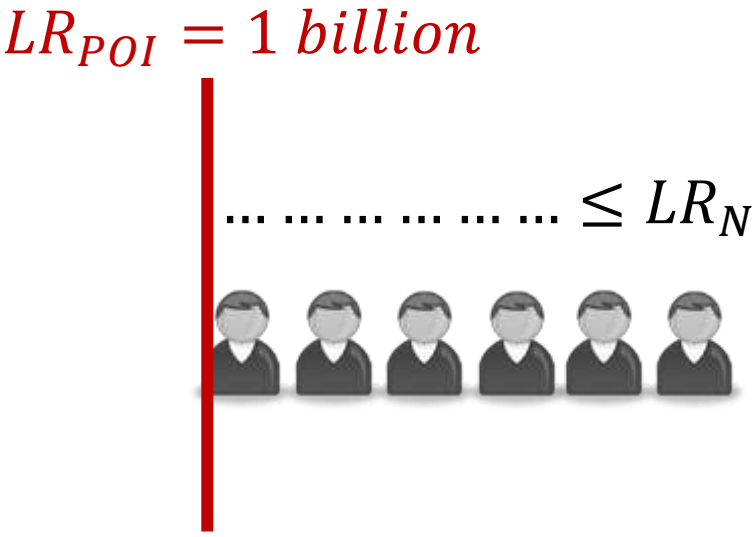
$= 1 - \delta$


$LR_{POI} = 1\ billion$
$LR_1 \leq LR_2 \leq LR_3 \leq \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \leq LR_N$
1
2
3
N
k individuals:
$LR_i \geq LR_{POI}$

Average of k LRs:

$$\frac{k}{N} LR_{POI} + \varepsilon$$

$$= 1 - \delta$$



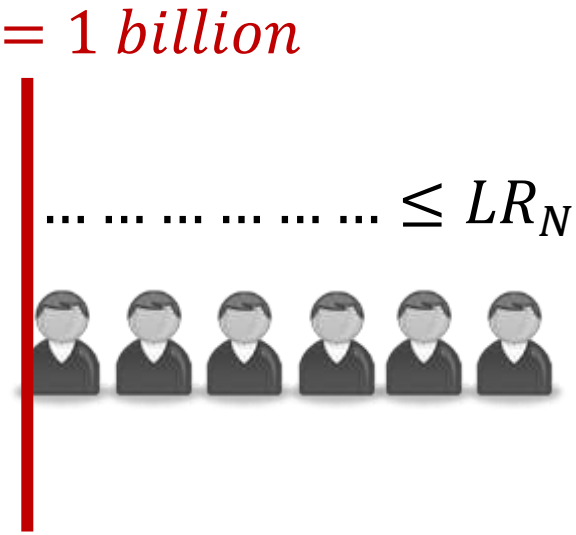
2) Information on the chance of adventitious matching is provided by the LR.

Average of k LRs:

$$\frac{k}{N} LR_{POI} + \varepsilon = 1 - \delta$$

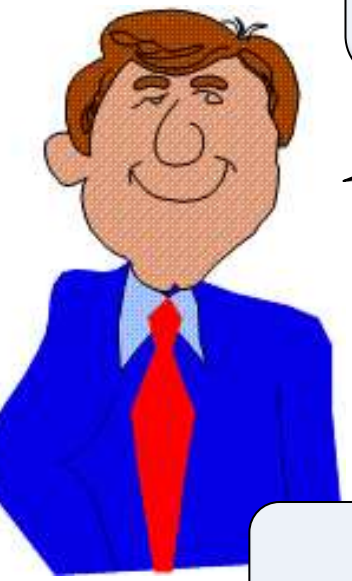
$$\frac{k}{N} = \frac{1 - \delta - \varepsilon}{LR_{POI}}$$

$$\frac{k}{N} \leq \frac{1}{LR_{POI}}$$



The proportion of randomly chosen non-donors expected to yield a likelihood ratio of LR_{POI} or larger is less than or equal to 1 in LR_{POI}

In court



What were the results of the comparison of the crime stain's DNA with the POI's DNA?

The DNA typing results of the crime stain are $1\ billion$ times more probable to have been obtained if the POI is the donor of this DNA than if an unknown person is the donor of this DNA.

How many other people would adventitiously match this profile?

The proportion of randomly chosen non-donors expected to yield a likelihood ratio of $1\ billion$ or larger is less than or equal to $1\ in\ 1\ billion$.

Three Main Points

1) The average value of the likelihood ratios (LRs) obtained for tests where H_d is true is 1.

2) Information on the chance of adventitious matching is provided by the LR.

3) This theory provides a framework for assessing the performance of interpretation models.

This theory also applies to mixtures.

3) This theory provides a framework for assessing the performance of interpretation models.

The model's performance can be evaluated using $H_d\ true$ tests.

1. Simulate random genotypes (these are the randomly chosen non-contributors)
2. Compare with the mixed DNA profile
3. Generate a LR for each with propositions

   H_p: The DNA came from *Random Person* and unknowns.

   H_d: The DNA came from unknowns.

3) This theory provides a framework for assessing the performance of interpretation models.

The model's performance can be evaluated using $H_d\ true$ tests.

→ Calculate average LR

→ Calculate the proportion of LRs greater than or equal to the known contributors' LRs

Conclusions

- Tests where H_p is true
- Tests where H_d is true:

Is the average LR close to 1?

Is the proportion of LRs greater than or equal to the known contributor's LR smaller than or equal to 1 over this LR?

Acknowledgements

US National Institute of Justice
grant 2011-DN-BX-K541

Thank you for your attention