Netherlands Forensic Institute
Ministry of Security and Justice

The NFI experience with the statistical evaluation of complex DNA mixtures.

Ate Kloosterman
Netherlands Forensic Institute[1] &
University of Amsterdam[2]

May 28 2014

a.kloosterman@nfi.minvenj.nl

DNA profile comparison

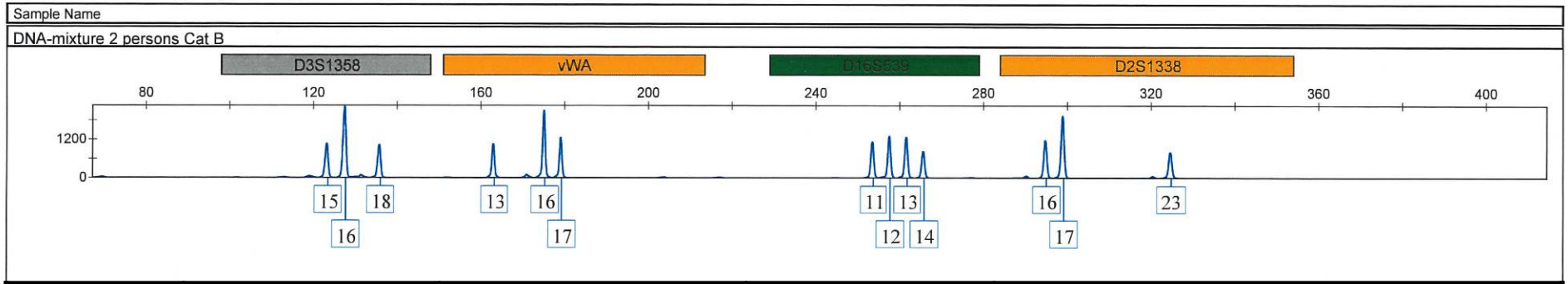
NFI approach: result of a DNA profile comparison is categorized to one of four categories

- A: Exclusion
- B: Match with statistical interpretation
- C: Match without statistical interpretation
- D: No exclusion or match; subdivided into:
  - D1: Cannot be excluded
  - D2: Inconclusive

Definition of a DNA match (categories B and C):

"it can be concluded, that all the alleles of the reference profile are also present in the defined DNA profile of the evidence".

Category B: match with statistical interpretation



	D3S1358	vWA	D16S539	D2S1338
Evidence	**15 16 18**	**13 16 17**	**11 12 13 14**	**16 17 23**
Suspect	**15 18**	**13 16**	**11 14**	**17 23**

- DNA profile suspect matches mixed DNA profile evidence
- The assumption that all alleles of all donors are present in the mixed DNA profile of the evidence can be safely made

DNA profile comparison and evidential value

Category B: Match with statistical interpretation

- Matching DNA profiles: determine whether the evidential value of the match can be evaluated statistically.
- Condition: the determination or assumption that all the alleles of all the donors are present in the defined DNA profile of the evidence
- In practice generally only applied in case of (deduced) single source DNA profiles: random match probability and complete mixed DNA profiles of two contributors (some cases three contributors)
- likelihood ratio method (preferably) or probability of inclusion/exclusion

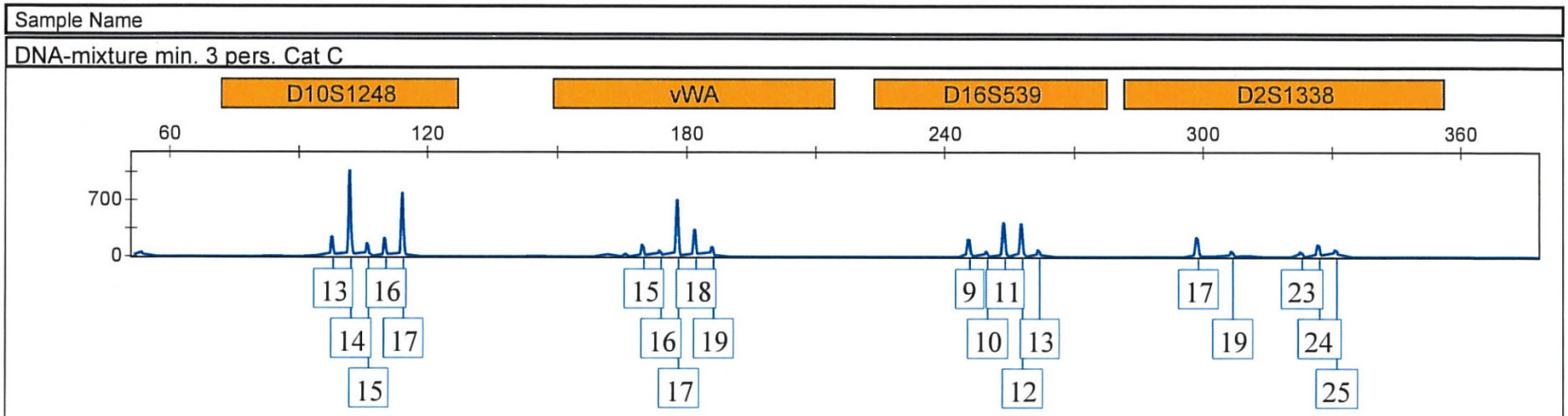
e.g. '*the result of the DNA profile comparison is x times more likely if the prosecution hypothesis is true, than if the defence hypothesis is true*'

2012: DNA match; no statistical interpretation

	D10S1248	vWA	D16S539	D2S1338
Evidence Consensus	**13 14 15 16 17**	**15 16 17 18 19**	**9 11 12 13**	**17 19 23 24**
Not in consensus (one observation)			**10**	**25**
Suspect	**13 16**	**15 17**	**9 12**	**17 24**

- DNA profile suspect matches mixed defined (consensus) DNA profile of the evidence
- It is uncertain if all alleles of all donors are present in the defined mixed profile of the evidence

2012: Person of Interest can not be excluded

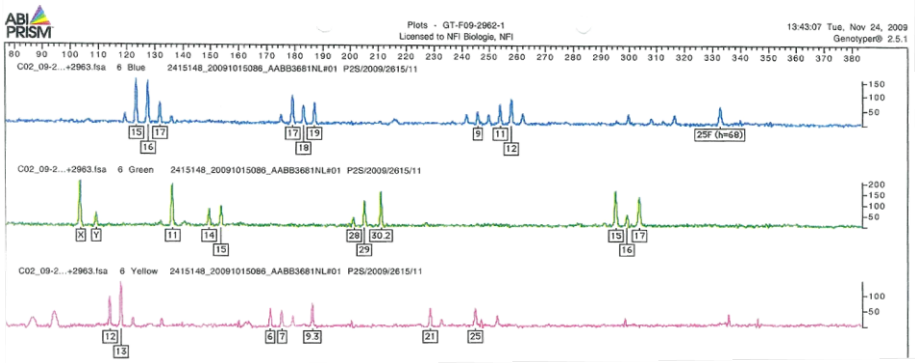


	D10S1248	vWA	D16S539	D2S1338
Evidence Consensus	**13 14 15 16 17**	**15 16 17 18 19**	**9 10 11 12 13**	**17 19 23 24 25**
Not in consensus (one observation)				**22 25**
Suspect	**15 17**	**18 19**	**10 12**	**22 24**

- DNA profile suspect does not match defined (consensus) mixed DNA profile of the evidence
- A relatively large number of alleles of the suspect occur in the defined (consensus) mixed DNA profile; questionable allele is observed in High MW locus (D2S1338) that is prone to allele drop-out

Inconclusive DNA analysis result

- Result supports neither a hypothesis about the presence of the person's DNA in the evidence nor one about its absence



DNA profile from highly degraded mixed DNA

Complex (LT) mixed DNA profile of a high, but unknown number of contributors

Over the years: increase in the number of complex DNA profiles

1. Increase in the number of case requests and samples
2. Absolute and relative increase in the number of non-classical DNA traces (touch evidence).
3. More requests for statistical evaluation of complex DNA profiles.

- Low DNA concentration
- NFI quantification method: Alu Quant real time PCR method
- Enhanced number of PCR cycles or enhanced detection (i.e. 9 kV CE injection)
- Three PCR replications (maximum volume as input for DNA analysis)

Match/non match with a complex forensic DNA profile

To report that two patterns match, without providing any scientifically valid estimate of the frequency with which such matches might occur by chance, is meaningless.
DNA “inclusions” cannot be interpreted by the CJS without statistical evaluation of the evidential value.

NATIONAL RESEARCH COUNCIL, DNA TECHNOLOGY IN FORENSIC SCIENCE 74-75 (1992).

It is also bad science to discard DNA profiling data and to believe that this is always the conservative approach as this is only true if the complex DNA profile does not show exclusionary information.

The statistical evaluation of complex mixed DNA profiles is a hot topic in court

Extra onderzoek naar DNA–spoor vermoorde Nicole van den Hurk, Jos de G. blijft vastzitten

Publicatie: donderdag 24 april 2014 - 14:44 | Auteur: **Sebastiaan Quekel**


Jos de G.

DEN BOSCH - Job Knoester, de advocaat van moordverdachte Jos van der G., krijgt van de rechter de kans nieuw onderzoek te laten uitvoeren naar het DNA-spoor op de vermoorde Eindhovense Nicole van den Hurk. Zijn verzoek om een second opinion wordt gehonoreerd. Dat werd donderdag duidelijk tijdens de eerste zitting in deze zaak in de rechtbank in Den Bosch.

De rechter meent dat er genoeg ernstige verdenkingen zijn tegen Jos de G.. De verdachte blijft daarom in voorarrest. De rechtszaak gaat 3 juli verder.

The statistical evaluation of complex mixed DNA profiles is a also hot topic in forensic science

Forensic Sci Int Genet. 2014 Mar;9:47-54. doi: 10.1016/j.fsigen.2013.10.011. Epub 2013 Oct 31.

Euroforgen-NoE collaborative exercise on LRmix to demonstrate standardization of the interpretation of complex DNA profiles.

Prieto L[1], Haned H[2], Mosquera A[3], Crespillo M[4], Alemañ M[5], Aler M[6], Alvarez F[1], Baeza-Richer C[7], Dominguez A[8], Doutremepuich C[9], Farfán MJ[10], Fenger-Grøn M[11], García-Ganivet JM[12], González-Moya E[13], Hombreiro L[14], Lareu MV[3], Martínez-Jarreta B[15], Meriqioli S[16], Milans Del Bosch P[17], Morling N[11], Muñoz-Nieto M[4], Ortega-González E[18], Pedrosa S[19], Pérez R[4], Solís C[1], Yurrebaso I[20], Gill P[21].

Forensic Sci Int Genet. 2013 Sep;7(5):555-63. doi: 10.1016/j.fsigen.2013.05.009. Epub 2013 Jul 11.

Evaluating forensic DNA profiles using peak heights, allowing for multiple donors, allelic dropout and stutters.

Puch-Solis R[1], Rodgers L, Mazumder A, Pope S, Evett I, Curran J, Balding D.

PLoS One. 2014 Mar 25;9(3):e92837. doi: 10.1371/journal.pone.0092837. eCollection 2014.

TrueAllele casework on Virginia DNA mixture evidence: computer and manual interpretation in 72 reported criminal cases.

Perlin MW[1], Dormer K[1], Hornyak J[1], Schiermeier-Wood L[2], Greenspoon S[2].

Proc Natl Acad Sci U S A. 2013 Jul 23;110(30):12241-6. doi: 10.1073/pnas.1219739110. Epub 2013 Jul 1.

Evaluation of mixed-source, low-template DNA profiles in forensic science.

Balding DJ.

Analysis of Forensic DNA Mixtures with Artefacts

R. G. Cowell, T. Graversen, S. Lauritzen, J. Mortera

The statistical evaluation of complex mixed DNA profiles is a also hot topic in forensic science

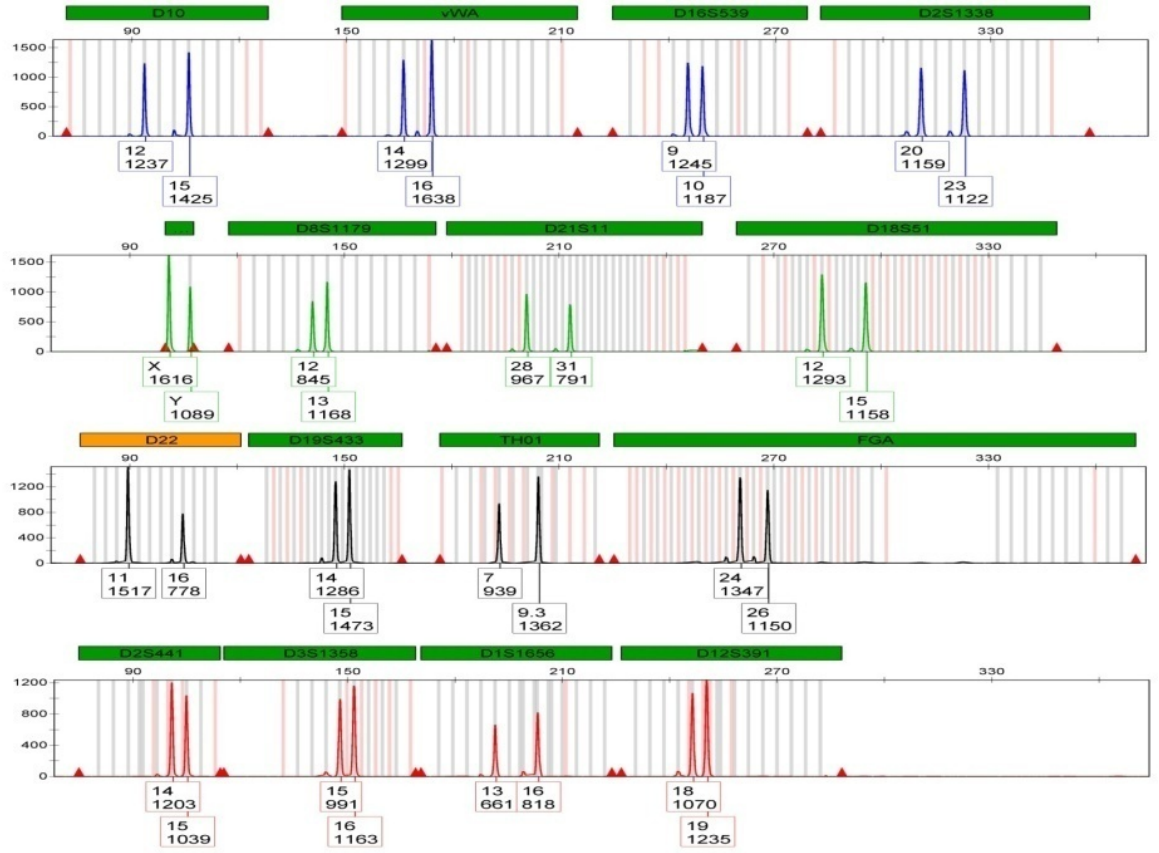
- Binary method (treats alleles as present or absent)
- Semi continuous methods treats alleles as present or absent but assigns a probability to the events of drop out and drop in.
- Full continuous models deal with the probability of drop out and other stochastic effects based on the intensity of the alleles visualized at a locus.

Implementing in forensic DNA analysis: diversion of models and methods

If our condition were truly happy, we would not seek diversion from it in order to make ourselves happy.

Blaise Pascal

In contrast with: standardization on an ESS forensic DNA profile



The technology enables world wide comparison of DNA data and worldwide DNA database searches.

NFI critical expertise

- Validation of and experience with state of the art commercially available DNA analysis systems (NGM)
- Large experience with the analysis of low level real casework samples
- Experience with semi continuous statistical models (LRmix)

The forensic challenge

ISO 17025 compliance for forensic methods in case work

- *Accepted by the scientific community*
- *Known and published error rates*
- *Validated under realistic case conditions*
- *Proficiency tests*

The NFI developed a simple semi continuous model

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

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

Exploratory data analysis for the interpretation of low template DNA mixtures

H. Haned [a,*], K. Slooten [a,b], P. Gill [c,d]

[a] Netherlands Forensic Institute, Department of Human Biological traces, The Hague, The Netherlands
[b] VU University Amsterdam, Amsterdam, The Netherlands
[c] Norwegian institute of Public Health, Oslo, Norway
[d] University of Oslo, Norway

The NFI developed a binary (semi continuous) model

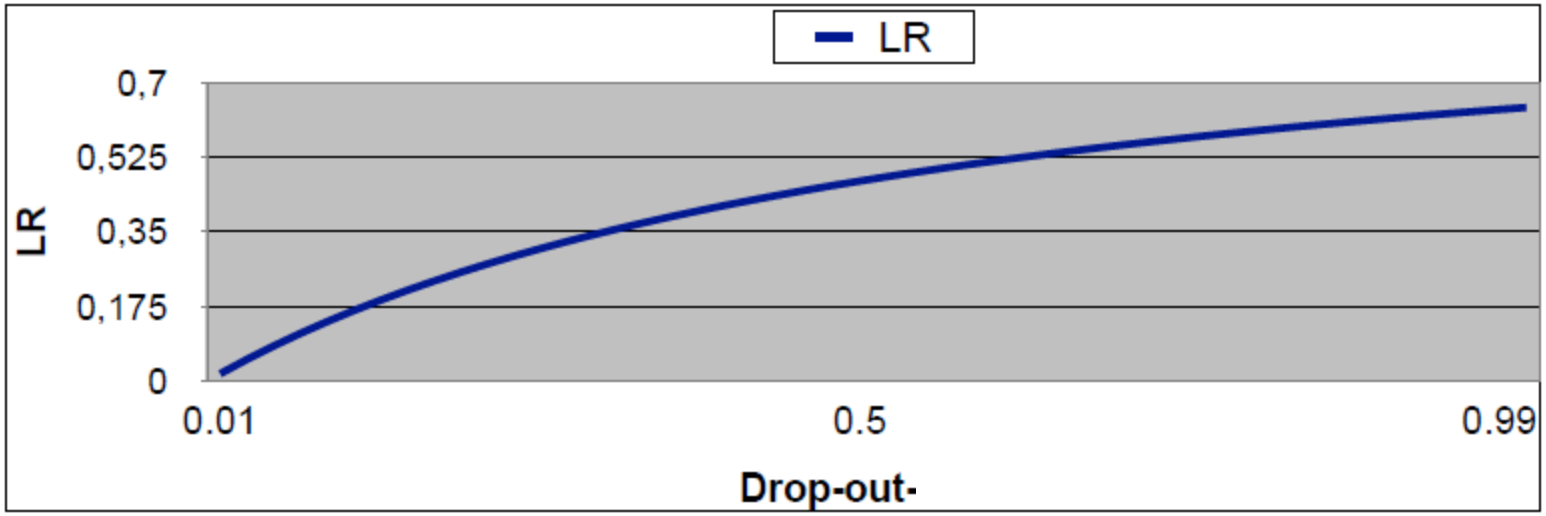
The NFI model

- alleles are present or not
- alleles of donors can dropout with probability D
- alleles can drop-in

- No use peak height information data
- No modulation of allele stutter (standard stutter settings)
- no DNA degradation
- no kinship
- We know this model has strengths (simple) and short comings (not all the available information is used), but it may be useful!

LR and drop-out probability (no drop-in)

data: 9,11,12 suspect: 9,10



The NFI model: an example four PCR replicates

D16 drop out 0.1, suspect 9,10

Data	Hp (Suspect+unknown)	Hd (2 unknowns)	LR
9,10,11,12	Unknown explains 11 and 12	2 unknowns explain 9,10,11 and 12	9.866
10,11,12	Unknown explains 11 and 12 Drop out of 9	2 unknowns explain 10,11 and 12	0.325
9,11,12	Unknown explains 11 and 12 Drop out of 10	2 unknowns explain 9,11 and 12	0.153
4 replicates 10,11,12 9,11,12 9,10,11,12 11,12	Unknown explains 11 and 12 Drop out of 9 Drop out of 10 - Drop out of 9 and 10	2 unknowns explain 9,10,11,12, Drop out of 9 Drop out of 10 - Drop out of 9 and 10	?

The NFI (semi continuous) model

Effect of drop-out probability
Effect of replicates

A low template mixture can give the same amount of evidence as a high template mixture!

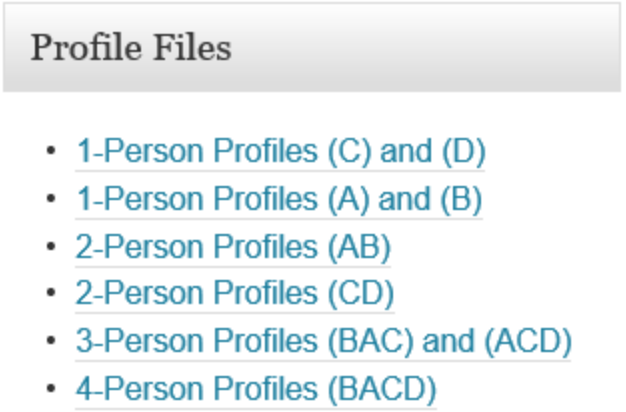
MINDSWITCH for our RO's

The mixed DNA profiles of interest

Examining the performance of existing probabilistic models in terms of:

1. the ability to discriminate between donors and non-donors
2. the stability of output data across reasonable settings of, for example, probabilities of drop-out and drop-in (calibration of the probabilistic model).
3. consideration of parameters such as reproducibility when replicates of the same mixture are analyzed.
4. Finally practical aspects such as time of analysis and ease of use are considered.

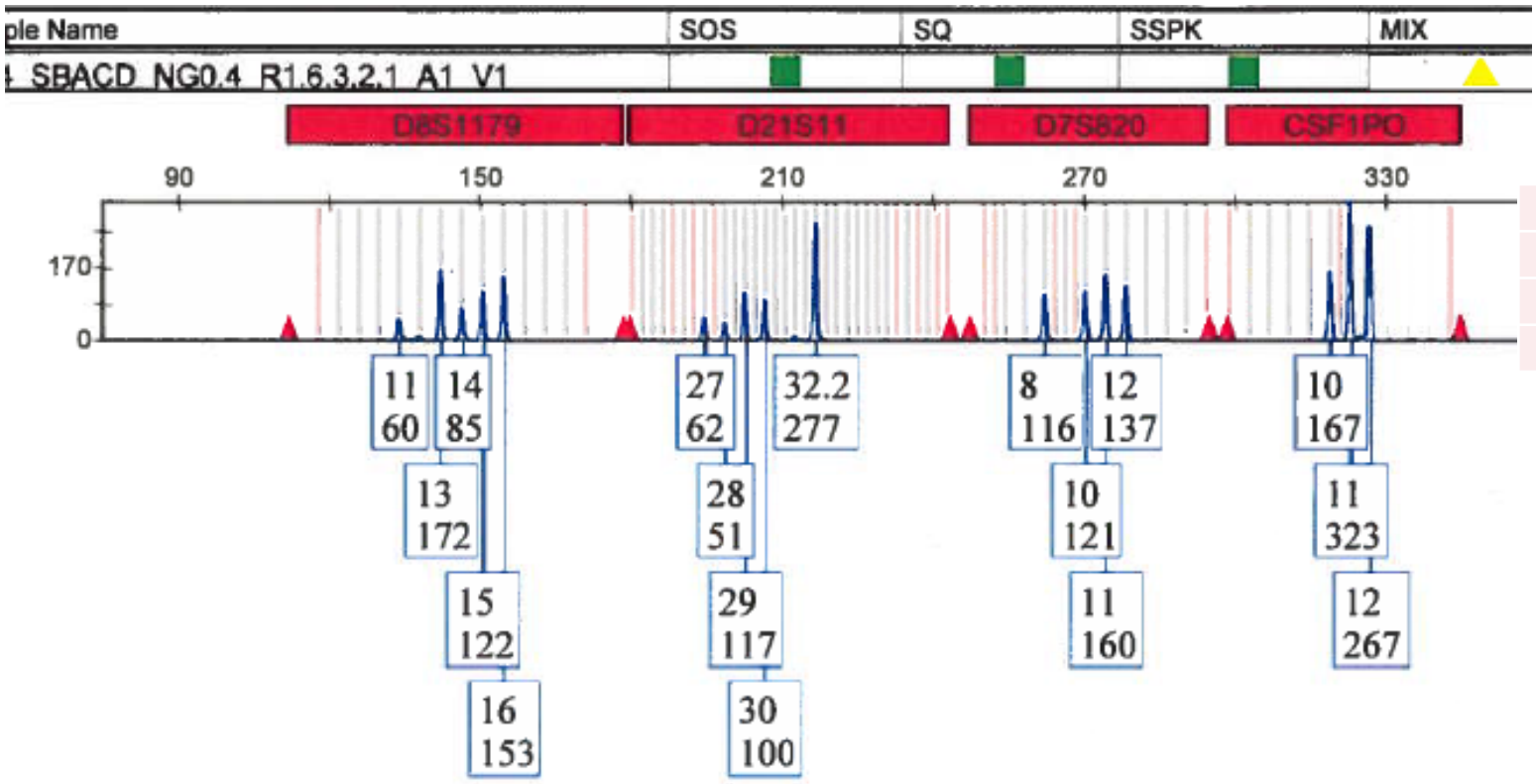
Webinar: the mixed DNA profiles of interest



- "open source" DNA profiles that have been made available by Robin Cotton from Boston University
- The mixtures include low-level contributors and two, three- and four-person contributors with differing ratios and allele sharing. Using the data from the mixed DNA-profiles we have investigated the performance of our model.

Example 4 person mixture

Four known donors: A, B, C and D (1.6:3:2:1)
Total DNA input: 0,4 ng



A	**84**	**pg**
B	**158**	**pg**
C	**105**	**pg**
D	**53**	**pg**

Identifiler DNA profiles of the BU samples

	A	B	C	D
D8S1179	13,16	11,13	14,15	13,16
D21S11	29,32.2	27,32.2	30,32.2	28,28
D7S820	8,11	11,11	10,12	8,12
CSF1PO	11,12	10,11	10,11	12,12
D3S1358	15,16	14,16	14,18	16,16
TH01	6,9	6,9.3	7,7	7,9.3
D13S317	11,11	11,13	11,12	12,13
D16S539	11,12	11,13	10,13	12,13
D2S1338	19,24	17,25	22,25	23,25
D19S433	15,15	14,15	12,14	13,13
vWA	18,19	15,18	15,16	15,19
TPOX	8,11	8,11	8,8	11,11
D18S51	13,14	13,17	16,18	14,20
Amel	XY	XX	XY	XY
D5S818	10,12	12,12	12,12	11,13
FGA	20,20	25,26	23,23	20,28

Input format of data: Genemapper Table output

ID_4_SBACD_NG0.4_R1.6,3,2,1_A1_V1_Identifiler_v1X

Sample File	Sample Name	D8S1179	Height	D21S11	Height	D7S820	Height		
CSF1PO	Height	D3S1358	Height	TH01	Height	D13S317	Height	D16S539	Height
D2S1338	Height	D19S433	Height	vWA	Height	TPOX	Height	D18S51	Height
AMEL	Height	D5S818	Height	FGA	Height				
ID_4_SBACD_NG0.4_R1.6,3,2,1_A1_V1.fsa	ID_4_SBACD_NG0.4_R1.6,3,2,1_A1_V1	11,	13,	14,	15,	16,	,	,	,
,	,	60,	172,	85,	122,	153,	,	,	,
,	27,	28,	29,	30,	32.2,	,	,	,	,
,	,	62,	51,	117,	100,	277,	,	,	,
,	8,	10,	11,	12,	,	,	,	,	,
,	,	116,	121,	160,	137,	,	,	,	,
,	,	,	10,	11,	12,	,	,	,	,
,	,	,	,	167,	323,	267,	,	,	,
,	,	,	14,	15,	16,	18,	,	,	,
,	,	,	,	114,	63,	150,	65,	,	,
,	,	,	,	,	6,	7,	9,	9.3,	,
,	,	,	,	,	132,	191,	57,	91,	,
,	,	,	,	,	,	11,	12,	13,	,
,	,	,	,	,	,	,	357,	145,	120,
,	,	,	,	,	,	10,	11,	12,	13,
,	,	,	,	,	,	104,	169,	129,	152,
,	,	,	,	,	,	,	19,	22,	24,
25,	,	,	,	,	,	86,	94,	130,	113,
,	,	,	,	,	,	12,	13,	14,	15,
,	,	,	,	,	,	66,	66,	137,	232,
,	,	,	,	,	,	15,	18,	19,	,
,	,	,	,	,	,	,	125,	143,	123,
,	,	,	,	,	,	8,	11,	,	,
,	,	,	,	,	,	406,	280,	,	,
,	,	,	,	,	,	,	13,	14,	16,
18,	,	,	,	,	,	202,	193,	108,	79,
,	,	,	,	,	X,	Y,	,	,	,
,	,	,	,	,	241,	154,	,	,	,
,	,	,	,	,	,	10,	12,	,	,
,	,	,	,	,	,	101,	317,	,	,
,	,	,	,	,	,	20,	23,	,	,
,	,	,	,	,	,	126,	97,	,	,
,	,	,	,	,	,	,	,		

The NFI model

LRmix tutorial, version 4.1

Hinda HANED
Netherlands Forensic Institute, The Hague, The Netherlands
May 2013

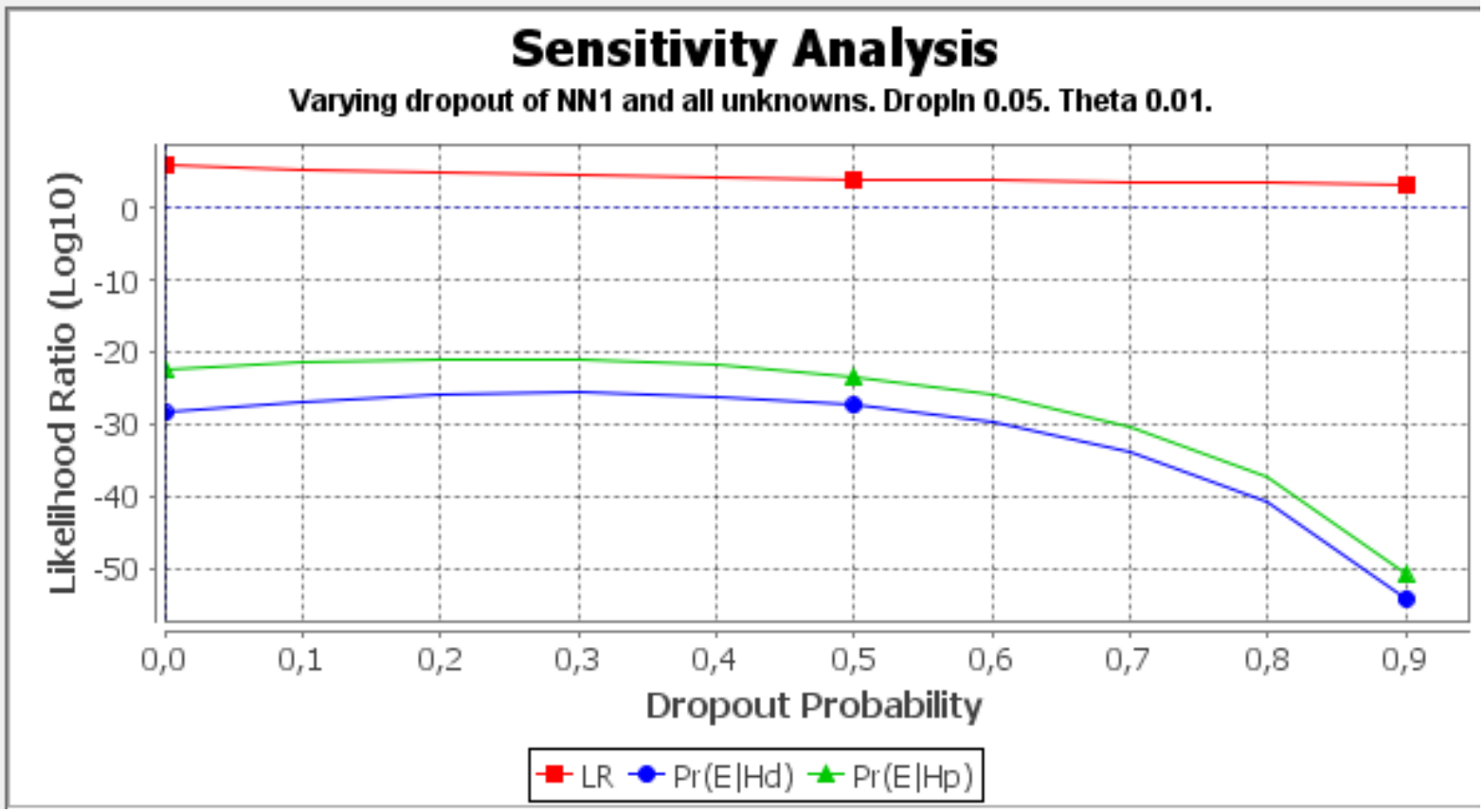
http://forensim.r-forge.r-project.org/misc/LRmix.pdf

The NFI model

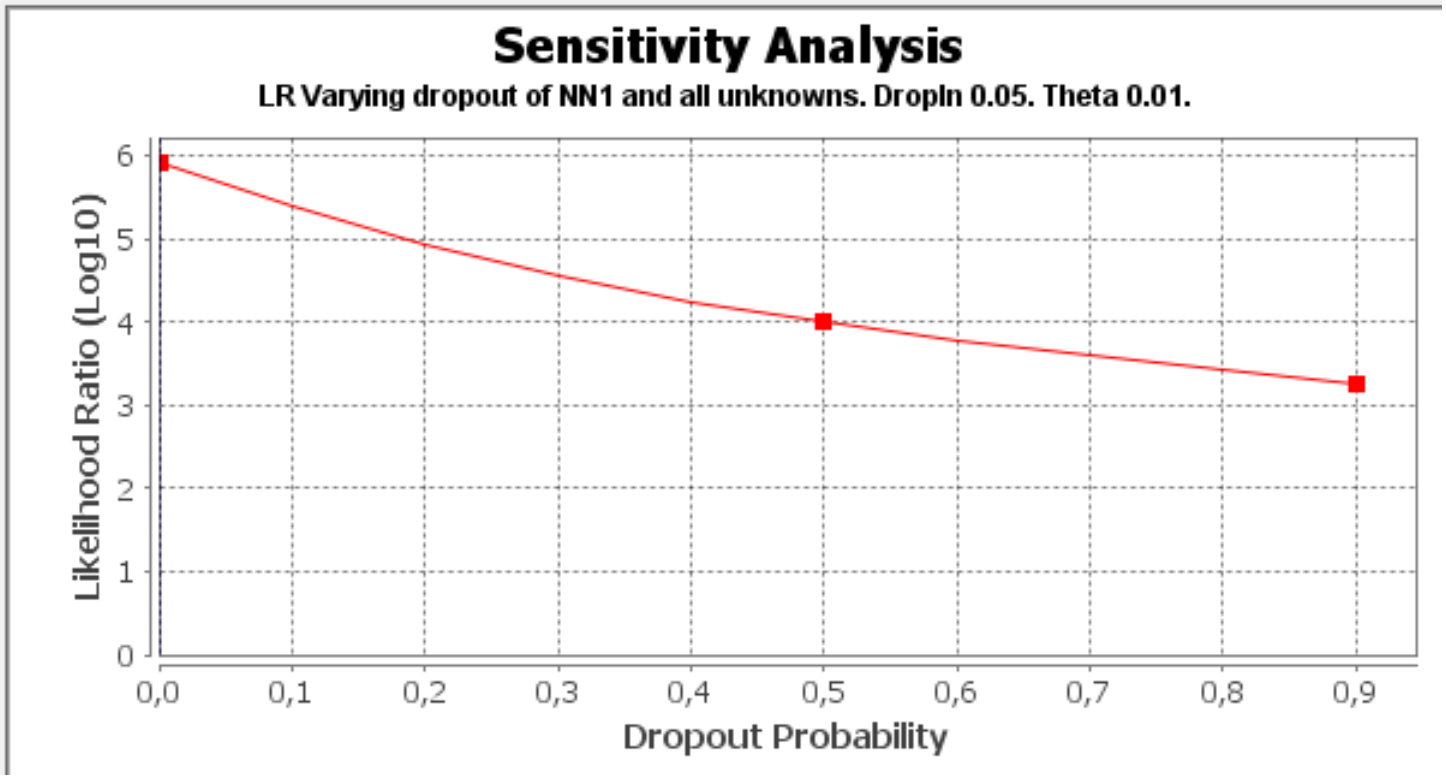
DEMO

Wrong number of actual contributors: two; not three

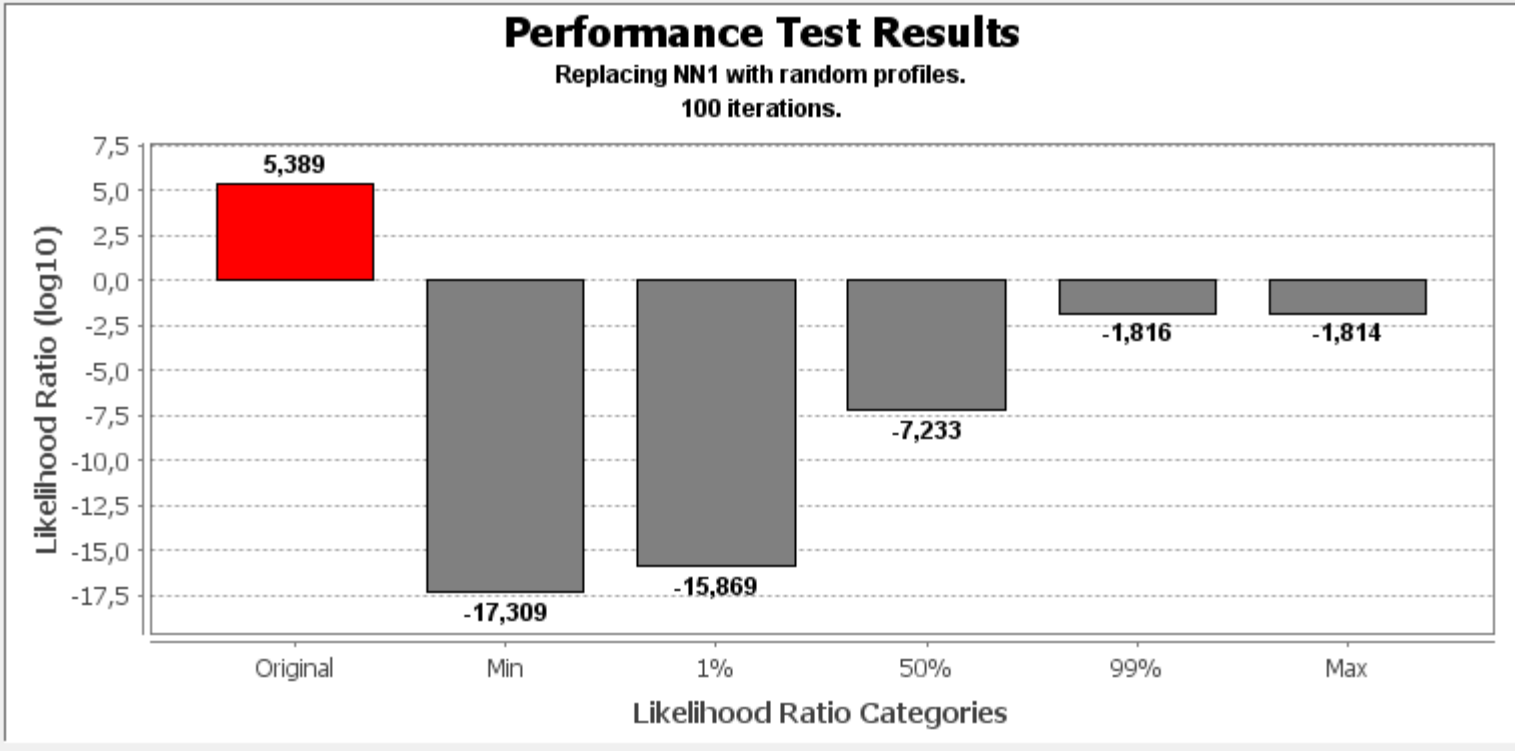
The NFI model



The NFI model



The NFI model



Identifiler DNA profiles of the BU samples

	A	B	C	D
D8S1179	3.73807	1.89406	2.04596	3.73807
D21S11	2.23084	8.30494	1.88078	0.97040
D7S820	1.82491	1.17489	1.62030	2.03096
CSF1PO	1.20197	1.42583	1.42583	1.05314
D3S1358	1.27660	2.08371	2.62520	1.05072
TH01	1.81562	1.02013	1.01472	1.12827
D13S317	1.38588	2.17541	1.60503	2.34612
D16S539	0.96580	1.42347	3.76649	1.41629
D2S1338	4.16796	0.36157	8.67447	0.36157
D19S433	1.20720	1.32525	1.98067	0.91796
vWA	4.18122	4.06806	0.38811	5.07756
TPOX	1.61543	1.61543	1.39012	1.61348
D18S51	3.15360	0.23667	4.21788	0.23226
Amel	5.87370	2.96616	2.96616	0.03172
D5S818				
FGA				
Total LR	245000	27	112000	0.44

A	84	pg
B	158	pg
C	105	pg
D	53	pg

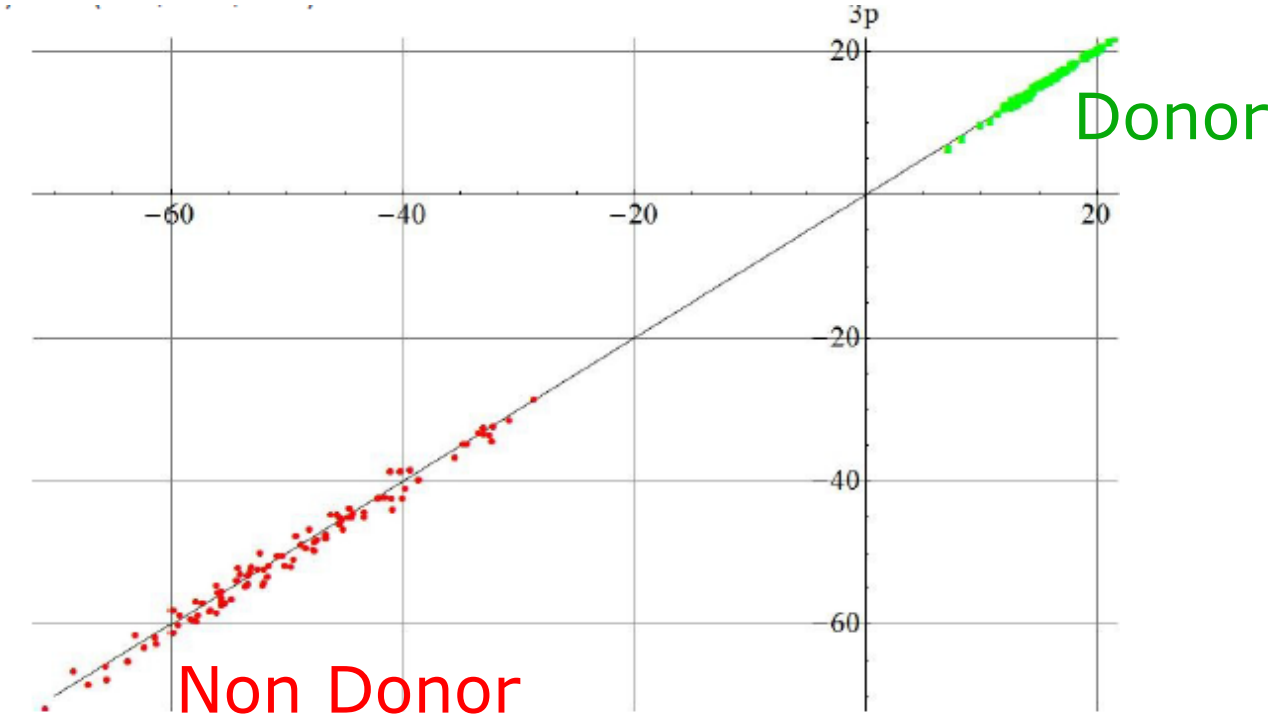
Scrutinise the model

Model uses parameters:

- Theta correction
- Allele frequencies
- Number of contributors
- Drop-out probability (per donor)
- Drop-in probability

What happens if the estimates of the parameters are incorrect: ROBUSTNESS

Wrong number of actual contributors: two; not three



Wrong number of drop out probability

Wrong number of contributors and wrong drop out rate

Scrutinise the model

The performance of the probabilistic model is examined in terms of:

1. the ability to discriminate between donors and non-donors
2. the stability of output data across reasonable settings of, for example, probabilities of drop-out and drop-in (calibration of the probabilistic model).
3. consideration of parameters such as reproducibility when replicates of the same mixture are analyzed.
4. Finally practical aspects such as time of analysis and ease of use are important.

ICFIS 2014 – Welcome to the Netherlands!

9th International Conference on Forensic Inference and Statistics

19-22 August, Leiden
www.icfis.org