When Drop-ins are Really Inconvenient:

Effect of Drop-in on False Positive and Rank-Order Likelihood Ratios Calculated by the Forensic Statistical Tool for a Mixture of Touch DNA

Clinton Hughes

The Legal Aid Society of New York City

Introduction

- I’m a lawyer, not a scientist
- Dr. Eli Shapiro: principle author and investigator
- Collaborated:
  - Directly impacts my casework right now
  - Highlight dangers of potential false positives in low template touch samples
  - Show the need for more access to information
  - Need for stronger emphasis for truly conservative analysis
- Slide 23 of 27

Defining the Problem

- Peter Gill: Defense at a disadvantage because of lack of access
- Example: FST is proprietary, lab is not sharing software or source code with the defense
- Courts have refused to order FST source code
- Solution: reverse engineering it, using public documentation
  - First, modified manual calculations
  - Second, computer science interns are writing code and creating a graphical user interface to replicate the results of the FST

What is the Forensic Statistical Tool (FST)?

- Semi-continuous LR model;
- Developed and used exclusively by the NYC Office of Chief Medical Examiner;
- Relies on pre-set drop-out and drop-in rates;
- Limited to two- and three-person mixtures;
- Cannot analyze four-person mixtures;
- Provides a single number, that is derived from prosecution and defense hypotheses that are dictated by the lab.
- The number is without context, other than a basic qualitative scale.

One trial court has rejected the black box, single score approach as not generally accepted. . . .

But the FST is, as a result, truly a “black box” – a program that cannot be used by defense experts with theories of the case different from the prosecution’s.

People v. Collins/Peaks, 2015 WL 4077176 (N.Y.Sup.) (Dwyer, J.)

. . . but other trial courts have found it admissible.

What is the PenB mixture?

- 60 pg (LCN) deducible mixture in Study 3E, a “clean touch” study in the FST False Positive Study.
- Cleaned with bleach, water and alcohol to remove DNA.
- Touched by three lab personnel, swabbed, DNA analysis and statistical analysis to determine the LR for known contributors.
- Subjected to “bulk run” against 546 morgue profiles, 700 NIST profiles, and LABTYPES databases.
- Over 480 bulk runs, Mitchell et al calculated a 0.003 False Positive Rate (excluding 10 LABTYPES hits) (highly disputed at hearing).
- PenB had 9 false positives (7 NIST, 1 morgue, 1 LABTYPES)

What is the PenB mixture (cont)?

- “JB,” male Caucasian profile from the NIST database, had the highest LR of those 9 false positives.
- JB scored a 156.79 as listed on the bulk run, or “Strong Support” (100-1000) Evett and Weir (1998).
- Using C++, programmer Kevin Ramdass obtained an LR of an 156.7903.
- Using Excel and DigDB, Arthur Speiser and I were eventually able to get an LR of 156.7902.

Why examine PenB?
FST Keeps Drop-in Rate Low

- Drop-in is defined as “contamination from an unknown source.” Butler (2010).
- Drop-in should weaken the strength of the evidence.
- Drop-in should be rare.
- Mitchell et al measured drop-in with pristine two- and three-person samples:

High Copy Number	
pC0	0.975
pC1	0.02
pC2	0.005

Low Copy Number	
pC0	0.96
pC1	0.035
pC2	0.005

Why Examine PenB?
FST counts stutter as drop-in

- Consistent with ISFG recommendation, the lab does not remove labels from peaks that are suspected stutter in mixtures;
- Stutter shouldn’t be reproducible from replicate to replicate.

Why examine the PenB Mixture? "Drop-ins" can lead to false positives

- In reality, drop-in rates for low template touch DNA mixtures amplified in triplicate can be quite high.
- A false positive profile can be derived from such drop-in, and without context, will mislead the fact finder as to the strength of the evidence.
- The fact finder is further misled by mixtures that are labelled three-person mixtures, when a more conservative analysis would be to call them as four-person, and therefore inconclusive.

JB: 156.79

3E_3PCPenB_60pg_3p_D_1000RR

Degraded Type: Not Degraded Deducible: Yes DNA pg 60

Result	ID	Asian	Black	Caucasian	Hispanic
	'D13	918909.31	7.48E+06	204687.25	835540.5
	'MB 3346	376811.47	7764.47	60.24	1348.34
	'JB	250360.7	598.97	156.79	1153.01
	'WT51507	26227.51	527.39	343	42.52
	'D27	800.75	475.84	9.38	45.86
	'WT51510	325.12	40.24	5.96	4.54
	'UT57317	90.02	2.54	0.03	0.07
	'ZT80147	62.65	5.02	0.27	0.77
	'WT51342	50.5	6.87	3.55	3.34
	'D5	20.35	104.8	0.64	3.76
	'GT37351	11.7	50.34	2.53	3.22
	'ZT80863	7.97	341.95	1.53	1.42
	'ZT80131	6.31	178.39	0.45	7.69
	'D21	6	165.03	17.38	4.08
	'WT51378	5.79	189.6	3.65	10.51
	'D1	3.57	0.2	1.80E-03	0.12

The Allele Chart for PenB

ded 37	39-Study_3E_3PC_Pen_B_60pg_a Esther071910_69N COO	8, 12, 13, 14, 15	28, 29, 30, 30.2	8, 9, 10, 11	9, 10, 12	14, 15, 16, 17, 18	6, 8, 9, 9.3	8, 9, 14	9, 10, 11, 12, 13	19, 20, 22	12, 13, 13.2, 14, 15.2	16, 17, 18	6, 8, 9, 11	12, 13, 14, 17, 18	X	11, 13	19, 21, 22, 23, 24, 25, 26
3 person	40-Study_3E_3PC_Pen_B_60pg_b Esther071910_69N COO	8, 10, 12, 13, 14, 15	28, 29, 29.2, 30	8, 9, 11	10, 12	14, 15, 16, 17, 18, 19	6, 8, 9, 9.3	8, 9, 11, 12	10, 11, 12, 13	19, 20, 22, 25	12, 13, 13.2, 14, 15.2	16, 17, 18	8, 9	12, 14, 17, 18	X	11, 13	19, 21, 22, 23, 25
	43-Study_3E_3PC_Pen_B_60pg_c Esther071910_69N COO	8, 13, 14	29, 30, 30.2, 31	8, 9	9, 10, 12	14, 15, 16, 17, 18, 19	6, 8, 9.3	8, 9, 11, 13	10, 12, 13	19, 27	12, 13, 13.2, 14, 15.2	16, 17	8, 9	12, 13, 14, 16, 17, 18	X	10, 11, 12, 13, 14	19, 21, 22, 23, 25, 26
	Composite* profile	8, 12, 13, 14, 15	28, 29, 30, 30.2, 31	8, 9, 11	9, 10, 12	14, 15, 16, 17, 18, 19	6, 8, 9, 9.3	8, 9, 11	10, 11, 12, 13	19, 20, 22	12, 13, 13.2, 14, 15.2	16, 17, 18	8, 9	12, 13, 14, 17, 18	X	11, 13	19, 21, 22, 23, 25, 26
	44-Study_3E_3PC_Pen_B_60pg_abc Esther071910_69N COO	8, 12, 13, 14, 15	28, 29, 30, 30.2, 31	8, 9, 11	9, 10, 12	14, 15, 16, 17, 18, 19	6, 8, 9, 9.3	8, 9, 11, 12, 13	9, 10, 11, 12, 13	19, 20, 22, 27	12, 13, 13.2, 14, 15.2	16, 17, 18	8, 9, 11	12, 13, 14, 16, 17, 18	X	11, 13, 14	19, 21, 22, 23, 25, 26
cannot be excluded	**Donor 21**	13 15	29 29.2	11	12	15 16	9	11 12	9 13	20 24	14 15.2	14 18	8 9	14 17	X	11	19 23
major component	**Donor 13**	8 14	29	8	9 10	14 18	6 8	8 9	10 12	19 22	12 14	16 17	8 9	12 14	X	11 13	22 25
cannot be excluded	**Donor 23**	10 15	28 29	10	12 14	15 16	6	12 13	11	19 25	14	14 16	8	13 14	X	11 12	23 24

- 75 distinct alleles in the mixture
- 23 distinct drop-in alleles
- The drop-in alleles tend to repeat and ‘three-peat’ across the replicates
- **7 drop-in alleles shared in JB’s profile**
- We will see how those drop-in alleles affect his LR

	D8S1179
Rep. 1	8,**12,**13,14,15
Rep. 2	8,10,**12,**13,14,15
Rep. 3	8,13,14
JB	**12,14**
Major (D13)	8,14
minor cbe	13,15
minor cbe	10,15

- JB's genotype of 12,14 is now fortuitously included in two of the three reps because of the drop-in allele 12;
- JB's 14 allele is in all three reps so his 12 appears to "drop-out" out of rep 3;
- JB's likelihood ratio of **1.71** is low because of the drop-out but tends to incriminate him;
- Our "reverse engineering" allows us to rank the LRs on a locus by locus basis. . .

D8S1179 (cont).

Of the 28 potential genotypes, twelve have LRs above 1.

JB's LR of 1.71 is inclusionary, but there are six genotypes above his and four even above the major. The "drop-in 12" allele leads to the highest LR for the locus, **20.82,** with an LR five times higher than the major's LR.

So we are able to examine JB's overall 15-locus LR on a locus by locus basis, to examine the effect of drop-in.

LRs Hi to Lo	
8,**12**	**20.82**
8,15	18.40
8,10	14.60
8,8	6.62
8,14 (major)	3.42
8,13	3.05
12,14 (JB)	**1.71**
14,15	1.51
12,13	1.49
13,15 (minor 1)	1.31
10,14	1.22
10,13	1.06

A more dramatic example of drop-in

	D7S820
Rep. 1	8,**9,**10,11
Rep. 2	8,**9,**11
Rep. 3	8,**9**
JB	**9,9**
Major (D13)	8,8
minor 1 cbe	11,11
minor 2 cbe	10,10

Dr. Shapiro noted that drop-in created JB's highest LR for the mixture, and also for the potential genotypes to this mixture.

JB as well as the three known contributors were all **homozygous** at the D7 locus. There is no issue of allele sharing, so we can see how the drop-in allele affects his LR.

The drop-in of the 9 allele in all of the replicates created JB's highest LR of **5.76** for the entire 15-locus mixture, and also for the potential genotypes to the D7 locus. . .

D7S820 (cont.)

Of the 15 possible genotypes, 7 have LRs above 1, all ranked higher than the two known minor contributors.

JB's profile is even higher than the deduced major profile

Dr. Shapiro's 'Bizarro World' Calculation: Without the 9 allele dropping into all three replicates, the overall 15-locus for JB would have been below 1.

Drop-in at one locus can lead to a false positive in the Strong Support category.

LRs Hi to Lo	
9,9 (JB)	**5.76**
8,8 (major)	5.00
8,9	2.79
9,11	2.23
9,10	2.06
8,11	1.96
8,10	1.80

Summer of 2014

Defense Lawyer:

> “So can you tell the jury the percentage of the population of New York City who FST would include in the mixture?

Testifying Criminalist:

> “No, I cannot.”

This is not good for anyone, particularly the jury.

Calculating the percentage of LRs above 1 for all non-contributors

D8S1179	
a,a	6.617509
a,b	14.59524
a,c	20.81642
a,d	3.046748
a,e	3.419212
a,f	18.40152
a,w	0.420912
b,b	0.000148
b,c	0.50037
b,d	1.061858
b,e	1.217779
b,f	0.440574
b,w	9.77E-05
c,c	0.063738
c,d	1.486298
c,e	1.708579
c,f	0.566191
c,w	0.013058
d,d	0.496814
d,e	0.289831
d,f	1.314008
d,w	0.0316
e,e	0.568451
e,f	1.510449
e,w	0.036157
f,f	0.056553
f,w	0.011586
w,w	1.21E-06

D21S11	
a,a	0.178728
a,b	1.48668
a,c	0.83888
a,d	1.032893
a,e	6.735722
a,f	6.175493
a,w	0.027306
b,b	1.054291
b,c	3.050258
b,d	0.555922
b,e	3.625008
b,f	3.502621
b,w	0.101509
c,c	0.002124
c,d	2.147969
c,e	13.42876
c,f	2.073577
c,w	0.000515
d,d	0.742687
d,e	2.517469
d,f	2.430509
d,w	0.071507
e,e	4.629768
e,f	15.86307
e,w	0.445763
f,f	0.416293
f,w	0.0636
w,w	3.9E-05

D7S820	
a,a	4.997265
a,b	2.78657
a,c	1.795386
a,d	1.963187
a,w	0.237497
b,b	5.761808
b,c	2.056575
b,d	2.233762
b,w	0.273832
c,c	0.146293
c,d	0.488827
c,w	0.024503
d,d	0.820958
d,w	0.073245
w,w	0.003044

CSF1PO	
a,a	3.98E+00
a,b	9.27E+00
a,c	8.98E+00
a,w	1.23E-01
b,b	2.74E+00
b,c	1.18E+00
b,w	1.43E-01
c,c	2.65E+00
c,w	1.38E-01
w,w	1.94E-03

D3S1358	
a,a	0.091195
a,b	1.080659
a,c	1.0988
a,d	0.951277
a,e	2.713627
a,f	10.8764
a,w	0.002354
b,b	0.050533
b,c	0.607056
b,d	0.525567
b,e	1.499053
b,f	5.983057
b,w	0.001305
c,c	0.051378
c,d	0.534387
c,e	1.524219
c,f	6.084438
c,w	0.001326
d,d	0.044508
d,e	1.319567
d,f	5.260004
d,w	0.001149
e,e	0.126214
e,f	15.11059
e,w	0.003259
f,f	0.001489
f,w	8.85E-05
w,w	9.74E-07

TH01	
a,a	1.447887
a,b	2.01768
a,c	1.330213
a,d	1.021723
a,w	0.065024
b,b	2.550031
b,c	2.25207
b,d	1.72622
b,w	0.11452
c,c	0.665469
c,d	1.135413
c,w	0.028026
d,d	1.205915
d,w	0.054157
w,w	0.003475

D13S317	
a,a	2.305935
a,b	1.965358
a,c	1.294566
a,d	1.003159
a,e	2.289528
a,f	2.849311
a,w	0.059753
b,b	3.723559
b,c	2.004855
b,d	1.58489
b,e	3.638487
b,f	4.543718
b,w	0.096488
c,c	0.110382
c,d	0.308974
c,e	0.73988
c,f	3.129836
c,w	0.021265
d,d	0.003562
d,e	0.548051
d,f	0.748467
d,w	0.005101
e,e	0.007248
e,f	1.70098
e,w	0.010381
f,f	0.008126
f,w	0.011638
w,w	6.61E-06

D16S539	
a,a	0.887265
a,b	2.638478
a,c	1.370859
a,d	1.648176
a,e	2.428087
a,f	3.229908
a,w	0.032063
b,b	0.031554
b,c	1.537736
b,d	0.335041
b,e	2.728775
b,f	3.631615
b,w	0.006208
c,c	0.515243
c,d	0.95574
c,e	1.417663
c,f	1.88569
c,w	0.01862
d,d	2.56E-05
d,e	1.70498
d,f	2.272971
d,w	2.74E-05
e,e	0.917766
e,f	3.340224
e,w	0.033166
f,f	1.222594
f,w	0.044181
w,w	8.2E-08

D2S1338	
a,a	5.318769
a,b	2.431305
a,c	8.03986
a,d	3.757156
a,e	9.168146
a,w	0.454092
b,b	0.526427
b,c	4.901872
b,d	2.689961
b,e	30.67855
b,w	0.34266
c,c	1.804246
c,d	9.048913
c,e	104.0847
c,w	0.248041
d,d	0.021893
d,e	9.376104
d,w	0.108649
e,e	0.049981
e,w	0.248041
w,w	6.08E-05

D19S433	
a,a	2.469175
a,b	1.855548
a,c	30.76023
a,d	1.579799
a,e	19.39464
a,w	0.063983
b,b	0.658513
b,c	8.351683
b,d	0.428862
b,e	5.265756
b,w	0.017064
c,c	11.16278
c,d	7.110456
c,e	87.28537
c,w	0.28926
d,d	0.558189
d,e	4.483198
d,w	0.014464
e,e	7.034367
e,w	0.182281
w,w	1.3E-08

vWA	
a,a	4.829997
a,b	0.829479
a,c	2.277589
a,w	0.307216
b,b	3.365241
b,c	1.566877
b,w	0.214049
c,c	0.340674
c,w	0.193871
w,w	6.04E-05

TPOX	
a,a	0.063367
a,b	1.179293
a,c	4.028661
a,d	0.046184
a,w	0.002262
b,b	1.182274
b,c	1.524624
b,d	0.483749
b,w	0.039206
c,c	4.044135
c,d	1.662106
c,w	0.134111
d,d	0.026203
d,w	0.000935
w,w	0.000303

D18S51	
a,a	0.887265
a,b	2.638478
a,c	1.370859
a,d	1.648176
a,e	2.428087
a,f	3.229908
a,w	0.032063
b,b	0.031554
b,c	1.537736
b,d	0.335041
b,e	2.728775
b,f	3.631615
b,w	0.006208
c,c	0.515243
c,d	0.95574
c,e	1.417663
c,f	1.88569
c,w	0.01862
d,d	2.56E-05
d,e	1.70498
d,f	2.272971
d,w	2.74E-05
e,e	0.917766
e,f	3.340224
e,w	0.033166
f,f	1.222594
f,w	0.044181
w,w	8.2E-08

D5S818	
a,a	0.003397
a,b	0.95538
a,c	0.013334
a,d	1.088576
a,e	0.072189
a,w	0.002582
b,b	3.55898
b,c	0.478584
b,d	1.047533
b,e	1.420934
b,w	0.20749
c,c	0.001743
c,d	0.546576
c,e	0.024255
c,w	0.001325
d,d	4.071298
d,e	1.613827
d,w	0.237359
e,e	0.00491
e,w	0.003732
w,w	2.45E-05

FGA	
a,a	0.42663
a,b	1.752915
a,c	1.911095
a,d	2.578676
a,e	0.169199
a,f	4.593344
a,g	6.546471
a,w	0.01977
b,b	0.246149
b,c	0.976153
b,d	1.315827
b,e	0.088661
b,f	2.341435
b,g	3.346926
b,w	0.011406
c,c	0.263455
c,d	1.434392
c,e	0.09626
c,f	2.552928
c,g	3.647004
c,w	0.012208
d,d	0.335666
d,e	0.128245
d,f	3.445416
d,g	4.914107
d,w	0.015555
e,e	2.34E-05
e,f	0.224
e,g	0.009315
e,w	0.000303
f,f	0.545657
f,g	8.748764
f,w	0.025285
g,g	0.002008
g,w	0.001554
w,w	4.78E-07

4 percent of the randomly-created Caucasian profiles

Forensic Science International: Genetics 9 (2014) 93–101

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Forensic Population Genetics – Original Research

Exact computation of the distribution of likelihood ratios with forensic applications

Guro Dørum [a,*], Øyvind Bleka [b], Peter Gill [b,c], Hinda Haned [d], Lars Snipen [a], Solve Sæbø [a], Thore Egeland [a,b]

- Probability of observing a LR as least as large as the suspect's LR.
- But the two known minors in this mixture were below JB's LR!

Is 157 really that strong?

- Peter Gill has questioned whether the qualitative scale for LRs associated with forensic science generally should be associated with LRs applied to DNA mixtures, where the risk of error is greater.

Why not analyze PenB Mixture as a four person mixture?

- OCME protocols allow for great latitude in calling a mixture a three-person mixture:

FORENSIC BIOLOGY PROTOCOLS FOR FORENSIC STR ANALYSIS

STR RESULTS INTERPRETATION		
DATE EFFECTIVE 02-02-2015	APPROVED BY NUCLEAR DNA TECHNICAL LEADER	PAGE 380 OF 508

b. Too many peaks labeled

i. Mixed HT-DNA samples that show seven or more labeled peaks (repeating or non-repeating) at two or more STR loci

ii. Mixed LT-DNA samples that show seven or more labeled peaks at two or more STR loci in the composite

314 FORENSIC SCIENCE

doi: 10.3325/cmj.2011.52.314

CMJ

Estimating the number of contributors to two-, three-, and four-person mixtures containing DNA in high template and low template amounts

Jaheida Perez, Adele A. Mitchell, Nubia Ducasse, Jeannie Tamariz, Theresa Caragine

Office of Chief Medical Examiner of the City of New York, The Department of Forensic Biology, New York, NY, USA

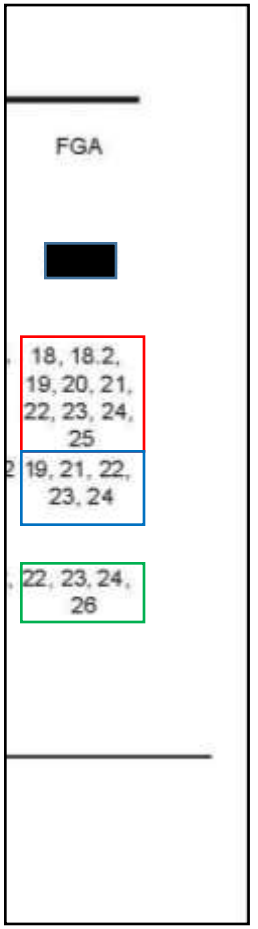
TABLE 1. The range, mean, and standard deviation of the number of different alleles expected in two-, three-, and four-person (p) mixtures

Mixture type	N	Maximum	Minimum	Mean	Standard deviation (SD)	Mean -2 SD	Mean +2 SD
2p	57	51	37	45.19	3.19	38.81	51.58
3p	105	66	47	57.23	3.68	49.86	64.59
4p	109	75	57	66.55	3.75	59.05	74.05

The maximum number of alleles observed for 109 4-person mixtures was 75, the same number as PenB!

PenB mixture resembles real casework on my desk right now

- Mixture 1: 9 in first run, 10 distinct alleles across three replicates at FGA; 80 distinct alleles; “3-person mixture”
- Mixture 2: 7 distinct alleles across three runs at two different loci; 65 distinct alleles (mean for four person in Perez was 66.55); “3 person mixture”
- Mixture 3: 10 distinct alleles across three replicates at one locus; 70 distinct alleles in the mixture; “3-person mixture”



Three-person limit may violate Daubert

I understand, and you're both relying on the Perez study. And what you've just done is to give me one possible explanation why there might not be four contributors. But that doesn't mean there weren't. It seems to me you have to be able to nail it pretty conclusively down to three contributors in order to use the test when you acknowledge that it's only valid for up to three contributors.

TRANSCRIPT OF CRIMINAL CAUSE FOR EVIDENTIARY HEARING
BEFORE THE HONORABLE BRIAN M. COGAN
UNITED STATES DISTRICT COURT JUDGE

United States Courthouse
Brooklyn, New York

Thursday, June 25, 2015
10:00 a.m.

Frye and the problem of the lab estimating both the number and controlling the defense hypothesis:

[T]hat there are anywhere from three to six alleles at each locus does not mean that four individuals' DNA is not present. But the "black box" nature of the FST prevents any defense attorney from informing the jury of the likelihood ratio, should the prosecution estimate of the number of contributors be incorrect. The jury will hear only one number: the one that is produced by "the program" as it assesses the prosecution hypothesis, and a dictated so-called defense hypothesis.

People v. Collins/Peaks, 2015 WL 4077176 (N.Y.Sup.) (Dwyer, J.)

Adding a fourth contributor only to the denominator

Accepted Manuscript

Title: The effect of varying the number of contributors on likelihood ratios for complex DNA mixtures

Author: Corina C.G. Benschop Hinda Haned Loes Jeurissen Peter D. Gill Titia Sijen

"That **we obtained lower LRs more often when an extra contributor was considered under Hd** [the defense hypothesis] is to be expected, as a random person may fit more easily with more unknowns, and therefore increases the likelihood of Hd and lowers the LR." S. 3.2.4

Thank you!

- Eli Shapiro
- Arthur Speiser
- Kevin Ramdass
- Khagay Nagdimov
- Nicolas Corpuz
- Nathan Adams
- Jessica Goldthwaite
- Susan Friedman
- Larry Mueller
- Peter Gill
- Mike Coble
- Bruce Budowle
- Kareem Belt
- Deborah Colson
- Kristen Santillo
- Celia Givens
- Tina Luongo
- Shilpy Goswami
- Richard Torres
- Allison Lewis
- Ralph Hughes
- Dan Krane
- Ranajit Chakraborty
- NIST