Defining the limits of forensic DNA profile interpretation: 
An assessment of the information content inherent in complex mixtures

Clare Marsden, Norah Rudin, Keith Inman, Kirk Lohmueller
What is the risk of error when interpreting complex DNA profiles?

- We used ground truth samples to establish the foundational limits when interpreting complex DNA profiles
- Error defined as drawing an incorrect inference with regard to whether a specific individual has contributed DNA to a profile
Complex mixtures
Complex mixtures

\[ LR = \frac{\text{Pr (Evidence} \mid \text{Suspect is contributor})}{\text{Pr (Evidence} \mid \text{Suspect is not the contributor})} \]

LR values expectation:

LR > 1 if suspect in the mixture.

LR < 1 if suspect not in the mixture.

Does this hold for complex mixtures?
Questions

- To what extent can LR reliably distinguish true-contributors (TC) and known non-contributors (KNC)?
  - Do true contributors always yield LR > 1 in complex mixtures?
  - Do known non-contributors give LR < 1 in complex mixtures?

CPI's as good as LR at distinguishing true-contributors from known non-contributors?
Simulations

Generating mixture set:

- C1
- C2
- C3
- C4
- C5
- KNC

True Contributor - in all mixtures

2 person mixture: C1 + C2
3 person mixture: C1 + C2 + C3

10,000 mixture sets created
Calculating the LR

\[
\text{LR} = \frac{\Pr(E \mid S)}{\Pr(E \mid S)}
\]

- True contributor
- Assumed contributor
- Unknown contributor
- Assumed contributor
- 2/3/4/5 person mixture
- 2/3/4/5 person mixture
Calculating the LR

How big are LR’s for true contributors?

\[ \text{LR} = \frac{\Pr(E \mid S)}{\Pr(S \mid E)} \]

Will true contributors always yield LR’s > 1?
Calculating the LR

\[ LR = \frac{Pr (E \mid S)}{Pr (E \mid \bar{S})} \]
Calculating the LR

LR = \frac{Pr(E|S)}{Pr(E|\bar{S})}

How small are LR’s for known non-contributors?

Will known non-contributors always yield LR’s < 1?
### Number of LR’s

<table>
<thead>
<tr>
<th># contributors to mixture</th>
<th>Numerator Hypothesis</th>
<th>Denominator Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 mixture</td>
<td>C1 + C2</td>
<td>C2 + 1 unk</td>
</tr>
<tr>
<td>2 mixture</td>
<td>C1 + 1 unk</td>
<td>2 unk</td>
</tr>
<tr>
<td>3 mixture</td>
<td>C1 + C2 + C3</td>
<td>C2 + C3 + 1 unk</td>
</tr>
<tr>
<td>3 mixture</td>
<td>C1 + C2 + 1 unk</td>
<td>C2 + 2 unk</td>
</tr>
<tr>
<td>3 mixture</td>
<td>C1 + 2 unk</td>
<td>3 unk</td>
</tr>
<tr>
<td>4 mixture</td>
<td>C1 + C2 + C3 + C4</td>
<td>C2 + C3 + C4 + 1 unk</td>
</tr>
<tr>
<td>4 mixture</td>
<td>C1 + C2 + C3 + 1 unk</td>
<td>C2 + C3 + 2 unk</td>
</tr>
<tr>
<td>4 mixture</td>
<td>C1 + C2 + 2 unk</td>
<td>C2 + 3 unk</td>
</tr>
<tr>
<td>4 mixture</td>
<td>C1 + 3 unk</td>
<td>4 unk</td>
</tr>
<tr>
<td>5 mixture</td>
<td>C1 + C2 + C3 + C4 + C5</td>
<td>C2 + C3 + C4 + C5 + 1 unk</td>
</tr>
<tr>
<td>5 mixture</td>
<td>C1 + C2 + C3 + C4 + 1 unk</td>
<td>C2 + C3 + C4 + 2 unk</td>
</tr>
<tr>
<td>5 mixture</td>
<td>C1 + C2 + C3 + 2 unk</td>
<td>C2 + C3 + 3 unk</td>
</tr>
<tr>
<td>5 mixture</td>
<td>C1 + C2 + 3 unk</td>
<td>C2 + 4 unk</td>
</tr>
<tr>
<td>5 mixture</td>
<td>C1 + 4 unk</td>
<td>5 unk</td>
</tr>
</tbody>
</table>

- 14 pairs (TC & KNC) of possible hypotheses
- 140,000 LR’s total each condition
Simulations

Generating LR’s:

Assumptions:

• No drop out
• No peak height information was used
Do true contributors always yield LR > 1

True Contributors

LR ➔

LR < 1: 7 of 140,000
5 person mixtures with 1 + unk in numerator
Do known non-contributors always yield LR > 1

4-5 person mixtures with 1 + unk in numerator
Can LR reliably distinguish true-contributors and known non-contributors?
Conclusions

- LR distinguish TC from KNC in mixtures containing up to 3 contributors.
- LR will distinguish most TC from KNC in mixtures of 4 or 5 individuals
  - Some low percentage of overlap
- LR generate a low false negative (TC < 1) and false positive (KNC > 1)
  - Primarily found in mixtures of 4 or 5 individuals