Complexity Thresholds and Exclusion Criteria

Catherine M. Grgicak

October 15, 2012
Nashville, TN
Do your reports contain a ‘concluding statement’ (i.e. included, excluded)

1. Yes
2. No
3. Don’t write reports

Data from 95 responses
ISHI Mixture Workshop (Oct 2012)
Is this suspect (yellow boxes) not-excluded?

<table>
<thead>
<tr>
<th></th>
<th>AT</th>
<th>30RFU</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST</td>
<td>150RFU</td>
<td></td>
</tr>
<tr>
<td>PHR</td>
<td>0.2 (&lt;500RFU)</td>
<td>0.5 (&gt;500RFU)</td>
</tr>
</tbody>
</table>

Major:Minor 4:1
Can the mixture shown in the previous slide be used for exclusion purposes?

1. Yes
2. No
3. I don’t know

Data from 89 responses
ISHI Mixture Workshop (Oct 2012)
Is the suspect (yellow boxes) included or excluded as a potential contributor to the mixture presented above?

1. Included
2. Excluded
3. Inconclusive
4. The mixture was uninterpretable

Data from 82 responses
ISHI Mixture Workshop (Oct 2012)
Is this suspect (yellow boxes) not-excluded?

<table>
<thead>
<tr>
<th></th>
<th>AT</th>
<th>30RFU</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST</td>
<td>150RFU</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PHR</th>
<th>0.2 (&lt;500RFU)</th>
<th>0.5 (&gt;500RFU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major:Minor</td>
<td>4:1</td>
<td></td>
</tr>
</tbody>
</table>
Can the mixture shown in the previous slide be used for exclusion purposes?

1. Yes
2. No
3. I don’t know

Data from 96 responses
ISHI Mixture Workshop (Oct 2012)
Is the suspect (yellow boxes) included or excluded as a potential contributor to the mixture presented above?

1. Included
2. Excluded
3. Inconclusive
4. The mixture was uninterpretable

Data from 98 responses
ISHI Mixture Workshop (Oct 2012)
### Statement of the Problem

DNA Commission of the International Society of Forensic Genetics: *Recommendations on the Interpretation of Mixtures*

<table>
<thead>
<tr>
<th>Interpretation Steps</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td>Identify the presence of a mixture</td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td>Designation of allelic peaks</td>
</tr>
<tr>
<td><strong>Step 3</strong></td>
<td>Identify the number of contributors in the mixture</td>
</tr>
<tr>
<td><strong>Step 4</strong></td>
<td>Estimation of the mixture proportion or ratio of the individuals contributing to the mixture</td>
</tr>
<tr>
<td><strong>Step 5</strong></td>
<td>Consideration of all possible genotype combinations</td>
</tr>
<tr>
<td><strong>Step 6</strong></td>
<td>Compare reference samples</td>
</tr>
</tbody>
</table>

Possible to consider “*complexity threshold*” before proceeding?

*Step “2.5”*

---

Is this suspect (yellow boxes) not-excluded?

19 discrepancies. Is this enough to exclude this suspect? Or is the mixture missing too much information?

Should we exclude suspect? Or is the comparison inconclusive due to the low-level and complexity?
Correct Inclusions vs Exclusions

- Created 10,000 mixtures,
  - 10,000 individuals who ought to have been excluded
  - 10,000 individuals who ought to have been included
  - Perturbing the mixtures with increasing levels of drop-out
  - Determined the proportion of false inclusions and false exclusions with varying levels of “allowed allelic discrepancies” (τ).

\[
\tau = 0, 1, \ldots, 30
\]

\[
\Pr(D) = 0, \quad \Pr(D) \approx 0.3
\]

\[
13, 14 \quad 13, 16 \quad 16, 16 \quad \text{etc.}
\]

\[
13, 17 \quad 14, 19 \quad 20, 20 \quad \text{etc.}
\]
- When Pr(D) = 0, get correct inclusion 100% of the time.
- With increasing levels of DO, correct inclusion rates decrease (i.e. you are more likely to exclude a standard who ought to have been included as a potential contributor.
- To alleviate this incorrect exclusion rate, allow for some allelic discrepancy (i.e. allow for some allelic drop-out to explain the inconsistency between standard and mixture).
- When Pr(D) = 0, get correct exclusion 100% of the time with < 6 discrepancies
- With increasing levels of DO, correct exclusion rates increase (i.e. you are more likely to exclude a standard who ought to have been excluded as a potential contributor)

<table>
<thead>
<tr>
<th>Alleles of Non-Contributor Standard</th>
<th>Non-Contributor Individual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locus 1</td>
<td>Locus 2</td>
</tr>
<tr>
<td>7, 11</td>
<td>8, 9</td>
</tr>
<tr>
<td><strong>Detected Mixture Alleles</strong> 7, 8, 9, 11</td>
<td>5, 8, 9, 14</td>
</tr>
<tr>
<td><strong>Discrepant Allele(s)</strong> 11</td>
<td></td>
</tr>
<tr>
<td><strong>Allowed Discrepancy</strong> τ₀</td>
<td>Excluded since δ=1 &gt; τ₀</td>
</tr>
<tr>
<td><strong>Allowed Discrepancy</strong> τ₁</td>
<td>Included since δ=1 ≤ τ₁</td>
</tr>
<tr>
<td><strong>Allowed Discrepancy</strong> τ₂</td>
<td>Included since δ=1 ≤ τ₂</td>
</tr>
<tr>
<td><strong>Detected Mixture Alleles</strong> 8, 9, 14</td>
<td>11, 13, 14</td>
</tr>
<tr>
<td><strong>Discrepant Allele(s)</strong> 7, 11</td>
<td></td>
</tr>
<tr>
<td><strong>Allowed Discrepancy</strong> τ₀</td>
<td>Excluded since δ=2 &gt; τ₀</td>
</tr>
<tr>
<td><strong>Allowed Discrepancy</strong> τ₁</td>
<td>Excluded since δ=2 &gt; τ₁</td>
</tr>
<tr>
<td><strong>Allowed Discrepancy</strong> τ₂</td>
<td>Included since δ=2 ≤ τ₂</td>
</tr>
<tr>
<td><strong>Detected Mixture Alleles</strong> 8, 9, 14</td>
<td>11, 13, 14</td>
</tr>
<tr>
<td><strong>Discrepant Allele(s)</strong> 7, 11</td>
<td></td>
</tr>
<tr>
<td><strong>Allowed Discrepancy</strong> τ₀</td>
<td>Excluded since δ=5 &gt; τ₀</td>
</tr>
<tr>
<td><strong>Allowed Discrepancy</strong> τ₁</td>
<td>Excluded since δ=5 &gt; τ₁</td>
</tr>
<tr>
<td><strong>Allowed Discrepancy</strong> τ₂</td>
<td>Excluded since δ=5 &gt; τ₂</td>
</tr>
</tbody>
</table>
Incorrect Inclusion v Correct Inclusion – R.O.C. Analysis

- Two-dimensional charts which plot the true positive versus the false positive rates for a given parameter or classifier.
- If the method is behaving perfectly then the false positive rate is 0 and the true positive rate is 1.

(0,1) = 0 incorrect inclusions and 100% correct inclusions
What percentage of the time would you be willing to falsely exclude a standard who in truth should be included?

1. 0%
2. 1%
3. 5%
4. 10%
5. 50%
6. 99%

Data from 88 responses
ISHI Mixture Workshop (Oct 2012)
What percentage of the time would you be willing to falsely include a standard who in truth should be excluded?

1. 0%
2. 1%
3. 5%
4. 10%
5. 50%
6. 99%

Data from 101 responses
ISHI Mixture Workshop (Oct 2012)
Increasing levels of DO, results in higher risk of incorrect inclusions and exclusions.
Complexity Threshold, Exclusion Criteria

- If the complexity criteria was set such that the lab does not want a correct inclusion rate < 85% and an incorrect inclusion rate > 1%, then mixtures with suspected Pr(D) > 0.3 should not be interpreted.

<table>
<thead>
<tr>
<th>Complexity Criteria:</th>
<th>Complexity Criteria:</th>
<th>Then suspected Pr(D) must be &lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Inclusion must be &gt; x%</td>
<td>Incorrect Inclusion must be &lt; y%</td>
<td>z</td>
</tr>
</tbody>
</table>

- 85% 1% 0.3
- 95% 0.1% 0.1
- 100% 0% 0

- In our lab, at AT of 30 RFU, Pr(D) ~ 0.3 when $\overline{H}$ ~ 70 RFU/allele or ~ 0.1 ng
- Pr(D) ~ 0.1 when $\overline{H}$ ~ 200 RFU or ~ 0.2 ng
- Pr(D) = 0 when $\overline{H}$ > 200 RFU or > 0.2 ng

<table>
<thead>
<tr>
<th>If the Complexity Criteria is</th>
<th>And suspected Pr(D) =</th>
<th>Then Exclusion Criteria (allelic discrepancies)</th>
<th>Overall Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 85% correct inclusion and &lt; 1% incorrect inclusion</td>
<td>0.3</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>
Is this suspect (yellow boxes) not-excluded?

Conclusion, this sample cannot be used for comparison purposes at a complexity threshold of 95% and 0.1%

A more lax complexity criterion such as 85% and 1%, would result in non-exclusion (τ=8) of the standard as a potential contributor. (Non-exclusion = included OR inconclusive)
Is this suspect (yellow boxes) not-excluded?

Conclusion, this profile cannot be utilized for comparison purposes
Conclusions

- Inherent risk of false inclusions and exclusions when using mixed, low-level samples for comparison purposes
- The level of drop-out, hence peak height, can be used to aid in determining whether the profile is suitable for comparison purposes
- This decision can be made BEFORE comparison to knowns
- R.O.C. analysis can be used as a tool to determine complexity and exclusion criteria
- If there is a need to “explain” why there are > 8 allelic discrepancies while still not excluding the standard, then with 2-person mixtures, there is a > 1% chance you are including a known that would have been excluded had you had a sufficient quantity of DNA and more stringent complexity guidelines.