ASSESSING THE QUALITY AND RELIABILITY OF THE DEA DRUG IDENTIFICATION PROCESS

Sandra E. Rodriguez-Cruz, Ph.D., ABC-F
DEA Senior Research Chemist
AAFS - February 16, 2017
Introduction

- **Background:**
  - DEA laboratory system (8 labs; > 270 chemists)
  - Tens of thousands reports per year
  - Produce accurate and scientifically-supported results

- **Objective:**
  - Quantitative assessment of the reliability of the overall laboratory process
  - Quality of laboratory results
  - Confidence (or uncertainty) of reported identifications
DEA Drug Identification Process:

Evidence Submission → Laboratory Report

I
COC, Barcoding, Storage...

II
Analytical Scheme

III
Report Preparation, T/A Review...

Established Analytical Techniques → Knowledge, Training, Expertise

Methods → Development/Validation
DEA LABORATORY ANALYTICAL SCHEME:

- Requires analysts to test, at minimum:
  - Two portions
  - Two different and independent techniques
  - Use negative controls
  - Use positive controls (traceable reference materials)

- SWGDRUG Recommendations

- ASTM E2329
  - Standard Practice for Identification of Seized Drugs
DEA Drug Identification Process:

- Where can errors occur?
- Phase I
  - Sample swapping, wrong barcoding, etc.
- Phase II
  - Analysis, sample swapping, contamination, etc.
- Phase III
  - Report preparation, dissemination, etc.
Uncertainty in Qualitative Analysis:

- Limited studies

- Past emphasis on quantitative analysis:
  - Measurement uncertainty

References:

DEA PTP HISTORICAL DATA:

- 2005-2016
- 4746 test results
- 2392 inter-laboratory (24-27 PT rounds/year)
- 2058 intra-laboratory
- 216 external
- 80 blind
CLASSIFICATION OF PT RESULTS:

- **All PTP Results**
  - **CS Present?**
    - **YES**
      - **CS Reported?**
        - **YES** True Positive
        - **NO** False Negative
    - **NO** False Positive
  - **NO** True Negative
**Calculating Response Rates:**

**TPR (Sensitivity):**

\[
TPR \ (sensitivity) = \frac{True \ Positives}{All \ Positives} = \frac{TP}{(TP + FN)}
\]

**TNR (Specificity):**

\[
TNR \ (specificity) = \frac{True \ Negatives}{All \ Negatives} = \frac{TN}{(TN + FP)}
\]

**FPR (Type I Error):**

\[
FPR \ (Type \ I \ error) = \frac{False \ Positives}{All \ Negatives} = \frac{FP}{(TN + FP)} = 1 - specificity
\]

**FNR (Type II Error):**

\[
FNR \ (Type \ II \ error) = \frac{False \ Negatives}{All \ Positives} = \frac{FN}{(TP + FN)} = 1 - sensitivity
\]
## RESULTS MATRIX:

<table>
<thead>
<tr>
<th></th>
<th>CS Reported</th>
<th></th>
<th></th>
<th>Total:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CS Present</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TPR (sensitivity)</td>
</tr>
<tr>
<td>YES</td>
<td>4285</td>
<td>4</td>
<td>4289</td>
<td>0.99907</td>
<td></td>
</tr>
<tr>
<td>NO</td>
<td>4</td>
<td>453</td>
<td>457</td>
<td>0.00875</td>
<td>FPR (type I error)</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td>4289</td>
<td>457</td>
<td>4746</td>
<td>0.00093</td>
<td></td>
</tr>
</tbody>
</table>

TPR (sensitivity) = 0.99907

FPR (type I error) = 0.00875

FNR (type II error) = 0.00093

TNR (specificity) = 0.99125
ABOUT THE FALSE RESULTS:

4 False Positives:
- Sample swapping
- Low-level secondary CS reported w/o fulfilling QA and documentation requirements
- 2 incorrect CS reported (LIMS)

4 False Negatives:
- Sample swapping
- Low concentration of target CS
- 2 cases of low-level secondary CS
**PRECISION AND ACCURACY:**

Precision = \( \frac{True \ Positives}{All \ Positives \ Results} \) = \( \frac{TP}{(TP + FP)} \)

\[
= \frac{4285}{(4285 + 4)} = 99.90\%
\]

Accuracy = \( \frac{All \ True \ Results}{All \ Results} \) = \( \frac{TP + TN}{(TP + FP + TN + FN)} \)

\[
= \frac{(4285 + 453)}{(4746)} = 99.83\%
\]
PT / RESPONSE RATES RESULTS:

- High sensitivity: 99.90%
- High specificity: 99.12%
- Low type I error rate: 0.87%
- Low type II error rate: 0.093%
- High precision: 99.90%
- High accuracy: 99.83%
USING BAYESIAN INFERENCE:

\[ P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \]

\[ P(CS|+) = \frac{P(+|CS) \cdot P(CS)}{P(+)} \]

\[ P(nCS|+) = \frac{P(+|nCS) \cdot P(nCS)}{P(+)} \]

CONFIDENCE IN THE POSITIVE ID:

\[
P(CS|+) = \frac{P(+|CS) \cdot P(CS)}{P(+|CS) \cdot P(CS) + P(+|nCS) \cdot P(nCS)}
\]

- True Positive Rate
- Prior probabilities
- False Positive Rate
- Posterior probability

- Probability CS is present, given a reported result
- Confidence in the positive identification result
Uncertainty in the Positive ID:

\[ P(nCS|+) = \frac{P(+|nCS) \cdot P(nCS)}{P(+|nCS) \cdot P(nCS) + P(+|CS) \cdot P(CS)} \]

- False Positive Rate
- Prior probabilities
- True Positive Rate
- Posterior probability

- Probability CS is not present, given a reported result
- Uncertainty in the positive identification result
ESTIMATING PRIOR PROBABILITIES:

- Population information
  - *Which population?*
- Historical data
- Prior knowledge
- Seizure circumstances
- Reasonable and supported assumptions

Candy Store

\[ P(CS) \rightarrow \text{low} \]
\[ P(nCS) \rightarrow \text{high} \]
ESTIMATING PRIOR PROBABILITIES:

- **No Information:**
  - \( P(CS) = P(nCS) = 0.50 \)

- **Prior Information:**
  - Statistics on laboratory submissions
  - Field testing results
  - Undercover purchase
  - Smuggling operations (clan lab, POE)
  - Identifying wrappings, markings, etc.
**NO POPULATION INFORMATION:**

- \( P(CS) = P(nCS) = 0.50 \)

Confidence =

\[
P(CS|+) = \frac{P(+|CS) \cdot P(CS)}{P(+|CS) \cdot P(CS) + P(+|nCS) \cdot P(nCS)}
\]

\[
P(CS|+) = \frac{(0.99907)}{(0.99907) + (0.00875)} = 0.9913 = 99.13\%
\]
No Population Information:

- $P(CS) = P(nCS) = 0.50$

Uncertainty =

$$P(nCS|+) = \frac{P(+|nCS) \cdot P(nCS)}{P(+|nCS) \cdot P(nCS) + P(+|CS) \cdot P(CS)}$$

$$P(nCS|+) = \frac{(0.00875)}{(0.00875) + (0.99907)} = 0.0086 = 0.86\%$$
ESTIMATING PRIOR PROBABILITIES:

- **No Information:**
  - $P(CS) = P(nCS) = 0.50$

- **Prior Information:**
  - Statistics on laboratory submissions
  - Field testing results
  - Undercover purchase
  - Smuggling operations (clan lab, POE)
  - Identifying wrappings, markings, etc.
# DEA Submissions & Reports:

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>Laboratory Results</th>
<th>CS (%)</th>
<th>NCS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CS</td>
<td>NCS</td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>37,115</td>
<td>32,779</td>
<td>4,336</td>
<td>88.32</td>
</tr>
<tr>
<td>1995</td>
<td>38,668</td>
<td>34,645</td>
<td>4,023</td>
<td>89.60</td>
</tr>
<tr>
<td>1996</td>
<td>43,662</td>
<td>38,836</td>
<td>4,826</td>
<td>88.95</td>
</tr>
<tr>
<td>1997</td>
<td>49,156</td>
<td>43,965</td>
<td>5,191</td>
<td>89.44</td>
</tr>
<tr>
<td>1998</td>
<td>55,946</td>
<td>49,919</td>
<td>6,027</td>
<td>89.23</td>
</tr>
<tr>
<td>1999</td>
<td>60,093</td>
<td>53,869</td>
<td>6,224</td>
<td>89.64</td>
</tr>
<tr>
<td>2000</td>
<td>64,608</td>
<td>57,840</td>
<td>6,768</td>
<td>89.52</td>
</tr>
<tr>
<td>2001</td>
<td>66,235</td>
<td>59,776</td>
<td>6,459</td>
<td>90.25</td>
</tr>
<tr>
<td>2002</td>
<td>64,504</td>
<td>58,065</td>
<td>6,439</td>
<td>90.02</td>
</tr>
<tr>
<td>2003</td>
<td>59,793</td>
<td>54,148</td>
<td>5,645</td>
<td>90.56</td>
</tr>
<tr>
<td>2004</td>
<td>56,709</td>
<td>50,973</td>
<td>5,736</td>
<td>89.89</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>596,489</strong></td>
<td><strong>534,815</strong></td>
<td><strong>61,674</strong></td>
<td><strong>88.20-90.96</strong></td>
</tr>
</tbody>
</table>

(95% Confidence Interval)
**POPULATION: DEA LAB SUBMISSIONS**

- $P(CS) = 0.90$
- $P(nCS) = 0.10$

Confidence = $P(CS|+) = \frac{P(+|CS) \cdot P(CS)}{P(+|CS) \cdot P(CS) + P(+|nCS) \cdot P(nCS)}$

$$P(CS|+) = \frac{(0.99907)(0.90)}{(0.99907)(0.90) + (0.00875)(0.10)}$$

$P(CS|+) = 0.99902 = 99.90\%$
POPULATION: DEA LAB SUBMISSIONS

- $P(CS) = 0.90$
- $P(nCS) = 0.10$

Uncertainty = $P(nCS|+) = \frac{P(+|nCS) \cdot P(nCS)}{P(+|nCS) \cdot P(nCS) + P(+|CS) \cdot P(CS)}$

$P(nCS|+) = \frac{(0.00875)(0.10)}{(0.00875)(0.10) + (0.99907)(0.90)}$

$P(nCS|+) = 0.00097 = 0.097\%$
## Confidence/Uncertainty:

<table>
<thead>
<tr>
<th>$P(CS)$</th>
<th>$P(nCS)$</th>
<th>Confidence (%)</th>
<th>Uncertainty (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>0.999</td>
<td>10.25</td>
<td>89.74</td>
</tr>
<tr>
<td>0.01</td>
<td>0.99</td>
<td>53.55</td>
<td>46.44</td>
</tr>
<tr>
<td>0.05</td>
<td>0.95</td>
<td>85.73</td>
<td>14.27</td>
</tr>
<tr>
<td>0.10</td>
<td>0.90</td>
<td>92.69</td>
<td>7.30</td>
</tr>
<tr>
<td>0.25</td>
<td>0.75</td>
<td>97.43</td>
<td>2.56</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
<td><strong>99.13</strong></td>
<td><strong>0.86</strong></td>
</tr>
<tr>
<td>0.75</td>
<td>0.25</td>
<td>99.70</td>
<td>0.29</td>
</tr>
<tr>
<td><strong>0.90</strong></td>
<td><strong>0.10</strong></td>
<td><strong>99.90</strong></td>
<td><strong>0.097</strong></td>
</tr>
<tr>
<td>0.95</td>
<td>0.05</td>
<td>99.95</td>
<td>0.046</td>
</tr>
<tr>
<td>0.99</td>
<td>0.01</td>
<td>99.99</td>
<td>0.009</td>
</tr>
<tr>
<td>0.999</td>
<td>0.001</td>
<td>99.99</td>
<td>0.001</td>
</tr>
</tbody>
</table>
**LIMITATIONS:**

- **Using PTP data:**
  - Not a ‘true’ representation of routine submissions?
  - Analyst “knows” it is a test
- ‘True’ sample identity not known
- No framework currently available
  - PTP data could be only data available to laboratories
- Prior probabilities (base rates) on the population
- Communicating approach to lay persons
CONCLUSIONS:

- PTP data provides means for assessing reliability of *overall laboratory drug identification process*.
- DEA laboratory ID process:
  - High sensitivity and specificity
  - Low type I and type II error rates
  - High accuracy and precision
- Bayesian & population assessment:
  - High confidence & low uncertainty
- Valuable assessment:
  - Improving laboratory testing & QA procedures
ACKNOWLEDGEMENTS:

- Ramona S. Montreuil  
  DEA Southeast Laboratory

- Mark E. McDonald, Richard P. Meyers, Scott R. Oulton, and Nelson A. Santos  
  DEA Office of Forensic Sciences

- Jason A. Bordelon, Jaclyn A. Iera, James V. Malone  
  DEA Southwest Laboratory

- Jack Mario  
  NY Suffolk County Crime Laboratory – retired

- Will Guthrie  
  NIST
QUESTIONS?

- Thank You!

sandra.e.rodriguez-cruz@usdoj.gov