Validation

John Butler
DNA Academy
University of Albany
June 13, 2005

Current Areas of NIST Research Effort

- Standard Information Resources (STRBase information, training materials/review articles, validation standardization, calibration datasets)
- Interlaboratory Studies (Real-time PCR, mixture interpretation)
- Resources for “Challenging Samples” (minisTRs for degraded DNA)
- Information on New Loci (Y-Chromosome, new STRs, SNPs)

Validation Project Purpose

- Review validation practices currently in use and available standards and guidelines (revised SWGDAM guidelines are too general)
- Help the community gain a better understanding of the validation process and how others have implemented validation in their labs so that validation in one’s own lab may be performed more quickly
- Attempt to define a minimum number of samples that could be recommended for various validation scenarios
- Help with establishing uniformity throughout the field to aid auditors in their inspections

Revised SWGDAM Validation Guidelines

(July 2004)

3. Internal Validation

…a total of at least 50 samples (some studies may not be necessary…)


National Institute of Justice
The Research, Development, and Evaluation Agency of the U.S. Department of Justice

Analytical Chemistry Application Review

June 15, 2005 issue of Analytical Chemistry

250 articles referenced
covering forensic DNA analysis during 2003-2004

Community Needs Training

- To better understand what validation entails and how it should be performed (why a particular data set is sufficient)
- Many labs already treat DNA as a “black box” and therefore simply want a “recipe” to follow
- People are currently driven by fear of auditors and courts rather than scientific reasoning
- Many different opinions exist and complete consensus is probably impossible

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Validation Definitions

ISO 17025

5.4.5.1 Validation is the confirmation by examination and the provision of objective evidence that the particular requirements for a specific intended use are fulfilled.

DAB Quality Assurance Standards for Forensic DNA Testing Laboratories

2 (ff) Validation is a process by which a procedure is evaluated to determine its efficacy and reliability for forensic casework analysis and includes:

- To demonstrate that a method is suitable for its intended purpose...

Pathway to Improved DNA Validation

- Collection of Current Philosophy on Validation
  - Community survey
  - Interviews
  - Literature summary
- Training
  - Auditors must be consistent in treatment of labs
- Providing Tools to Enable Improved Validation
  - Sample set(s)
  - Workbook – provide specific examples
  - Standard report form – documentation standardization
- Collection of Validation Data from Labs
  - NIST-funded labs to submit data to STRBase validation website

Contacting the Community

- Validation Standardization Questionnaire handed out at NIST DNA Grantees meeting (June 28-30, 2004)
- Emails sent to >200 scientists (July-Aug 2004)
  - Attendees from the NIST DNA Grantees meeting
  - Participants in NIST interlaboratory studies
  - Contacts through STRBase website
- Responses from 52 scientists were compiled
  - Covering 27 states + Puerto Rico, 4 companies, 2 outside US
- Specific interviews were conducted to gain perspectives from a small lab, a large lab, a private lab, and court testimony experience

Representative Labs Interviewed

- Montgomery County Crime Lab – small lab, 3 analysts, ~180 cases/year; using PP16 and ABI 310
- Orchid Cellmark – private contract lab, 40 analysts and technicians, ~5,000 cases/year; Profiler Plus/COFiler and Identifier with ABI 310 and ABI 3100; extensive court experience
- AFDL – large federal lab, ~120 analysts/technicians, remains identification rather than strictly forensic cases, >1,000 cases/year (mtDNA & STRs); Profiler Plus/COFiler and PP16 with ABI 377 and ABI 3100

Information from interviews is included in the written report of this project...
Review of Survey Questions

- What is validation?
- How do you know when you are finished validating a kit, instrument, software, or procedure?
- What steps are needed in internal validation and how many samples should be run at a minimum?
- How many total samples do you think it takes to internally "validate" a new forensic kit?
- How many different sets of samples are needed? Over what time period?
- Where do you look for guidance currently in terms of validation?
- What are some kits, software, instruments that you are considering for validation in the next year?
- How are validation, training, and proficiency testing related to one another?
- Do you think that the process of validation can be standardized?
- If a standard protocol or set of guidelines existed for validation, would you use it?
- If a standard set of samples existed for performing validation testing, would you use them?

Used to help define specific examples ...
Survey Summary for Recommended Precision Studies

A few of the responses:

• "100 allelic ladder injections"

• "1 allelic ladder with 10 injections"

• "Depends upon the system being tested. For a databanking system, 50-100 runs of 50-100 specimens. Again, stats tell you when you've processed enough specimens to understand the system."

• "Minimum: Run one sample at least 8 times.
Recommended: Run at least two samples plus allelic ladder at least 8 times." (24 sample-runs)

Survey Summary for Recommended Sensitivity Studies

Need to run samples that challenge interpretation at high DNA and low DNA concentrations—e.g., 10 ng and <0.2 ng

Most responses involve <10 samples with 10 ng to 30 pg range

Survey Summary for Recommended Mixture Studies

A few of the responses:

• 5 different 2-person mixtures
• 50 amplifications from at least 10 different mixtures
• 1 set of samples (ranging from 1:10 to 1:10:1)

Survey Summary for Recommended Non-Human Cases

A few of the responses:

• "10-20 food animals, companion animals, local wildlife, ferrets"

• "I don't believe this is necessary in internal validation if external results are published. This would not be expected to vary in different analysts' hands."

• "I've trusted system manufacturers to handle this. Should I have?"

• "Minimum: Include information from developmental studies. If performing developmental studies, include at least bacterial and yeast/fungal example, plus mammalian and non-mammalian examples."

Survey Summary for Recommended Non-Probative Cases

A few of the responses:

• Most responses were between 5-10 cases (range 3-25)

• "More important than the number of cases is the range of forensic samples that are typed during validation."

• "Complete cases are not required to test a system. Recommended: Run at least 8 mock non-probative samples. Note: Non-probative samples are not guaranteed to provide complete profiles. They are needed only to show that false results are not generated. Lack of results or incomplete results do not affect the validity of a validation."

Survey Summary for Recommended Numbers of Samples to Determine Heterozygote Peak Height Ratios and Stutter Values

A few of the responses:

• "100" samples

• "Minimum: 100 amplifications from at least 10 different sets of samples (ranging from 1:10 to 10:1)"
Where do you look for guidance currently in validation?

- SWGDAM
- DAB standards and ISO 17025
- Other scientists
- Literature publications
- Presentations at meetings
- Promega’s validation guide
- FBI studies and publications
- NIST studies and publications
- Previous scientific training
- Common sense

Validation Standardization Questionnaire (conducted June-August 2004)

Validation Section of the DNA Advisory Board Standards

STANDARD 8.1 The laboratory shall use validated methods and procedures for forensic casework analyses (DNA analyses).

8.1.1 Developmental validation that is conducted shall be appropriately documented.

8.1.3 Internal validation shall be performed and documented by the laboratory.

Revised SWGDAM Validation Guidelines
(July 2004)

3. Internal Validation
...a total of at least 50 samples (some studies may not be necessary...)

A Thoughtful Comment from One Interviewee

Before a set of validation experiments is performed...

- The question should be asked “Do we already know the answer to this question from the literature or a previous study performed in-house?”
- If the answer is “yes” and we document how we know this answer, then there is no need to perform that set of validation experiments.

A good example of this scenario is non-human DNA studies.
Common Perceptions of Validation

The goal is not to experience every possible scenario during validation...

"You cannot mimic casework because every case is different."

Many labs are examining far too many samples in validation and thus delaying application of casework and contributing to backlogs...

Significant time is required to perform studies

How an Assay Evolves

NIJ-funded project or company efforts

Research

Development

Optimization

Pre-Validation

Validation

Performance Check (Kit QC or Following Instrument Repair)

Performed by manufacturer

Performed by forensic lab

Implementation

Writing SOP, Training Others and Going "On-Line"

Steps Surrounding "Validation" in a Forensic Lab

Effort to Bring a Procedure "On-Line"

This is what takes the time...

Installation – purchase of equipment, ordering supplies, setting up in lab

Learning – efforts made to understand technique and gain experience; troubleshooting; can take place through direct experience in the lab or vicariously through the literature or hearing talks at meetings

Validation of Analytical Procedure – tests conducted in one’s lab to verify range of reliability and reproducibility for procedure

SOP Development – creating interpretation guidelines based on lab experience

QC of Materials – performance check of newly received reagents

Training – passing information on to others in the lab

Qualifying Test – demonstrating knowledge of procedure enabling start of casework

Proficiency Testing – verifying that trained analysts are performing procedure properly over time

A Comment on Minimum Numbers of Samples for Validation Studies

Impact of Number of Experiments on Capturing Variability in a Population of Data

From The HitchHiker’s Guide to the Galaxy

http://www.bbc.co.uk/dna/h2g2/

The Answer to the Ultimate Question Of Life, The Universe, And Everything

(and the Minimum Number of Samples for Internal Validation?)

42 + 8 = 50 (SWGDAM Revised Validation Guidelines)
Survey Summary of Planned Near-term “Validation”

**Commercial Kits**
- DNA IQ
- Qiagen
- Biomek 2000
- DNA Quant
- Quantifier

**STR Amp Kits**
- Identifier
- PowerPlex Y
- Yfiler
- PowerPlex 16
- ProPlus/COFiler reduced volume

**Software**
- GeneMapperID
- GentaScan
- Genotype NT
- TrueAllele
- SQL*LIMS and Forensic Solution

**Analysis Instruments**
- ABI 3100 Avant
- ABI 3100
- FMBIO III+
- MegaBACE

For RT-PCR
- ABI 7000
- Stratagene RT-PCR

---

**Internal Validation**

**Example: PowerPlex 16**
- Switch from ProfilerPlus/COFiler kits to PowerPlex 16
- Retaining same instrument platform of ABI 310

**Recommendations:**
- Concordance study (somewhat, but better to review literature to see impact across a larger number of samples and which loci would be expected to exhibit allele dropout e.g., DSS818)
- Stutter quantities, heterozygote peak height ratio
- Some sensitivity studies and mixture ratios
- Do not need precision studies to evaluate instrument reproducibility

---

**Internal Validation**

**Example: ABI 3100 Avant**
- Evaluation of a new ABI 3100 Avant when a laboratory already has experience with ABI 310
- STR kits used in lab will remain the same

**Recommendations:**
- Precision studies to evaluate instrument reproducibility
- Sensitivity studies
- Do not need new stutter, mixture ratio, peak height ratio, etc. (these relate to dynamics of the kit used)

---

**New Validation Homepage on STRBase**

http://www.cstl.nist.gov/biotech/strbase/validation.htm

---

**Validation Summary Sheet for PowerPlex Y**

- 1269 TOTAL SAMPLES EXAMINED
- 205 amounts (1/1.25/1.5/1.75/2 mM Mg) x 4 quantities (1/0.5/0.25/0.13 ng DNA) Magnesium titration
- 205 amounts (0.5x/0.75x/1x/1.5x/2x) x 4 quantities (1/0.5/0.25/0.13 ng DNA) Primer pair titration
- 205 amounts (1.38/2.06/2.75/3.44/4.13 U) x 4 quantities (1/0.5/0.25/0.13 ng DNA) TaqGold polymerase titration
- 102 females x 1 titration series (0-500 ng female DNA) x 5 amounts each Male-specificity
- 66 components of SRM 2395 NIST SRM 2424 animals Non-Human
- 505 volumes (50/25/15/12.5/6.25) x [5 amounts + 5 concentrations] Reaction volume
- 255 labs x 5 temperatures (54/58/60/62/64) x 1 sample Annealing Temperature
- 805 cycles (28/27/26/25/24) x 8 punch sizes x 2 samples Cycling Parameters
- 412 males used Stutter
- 52 males used with 52 samples Non-Probative Cases
- 66 lab x 2 series x 6 amounts (1/0.5/0.25/0.125/0.06/0.03) Sensitivity
- 6 labs x 2 M/M mixtures series x 11 ratios (1:0, 1:9, 1:1, 1:10, 1:100, 1:300, 1:1000, 0.5:300, 0.25:300, 0.125:300) Mixture Ratio (male:male)
- 6 labs x 2 M/F mixture series x 11 ratios (1:0, 1:1, 1:10, 1:100, 1:300, 1:1000, 0.5:300, 0.25:300, 0.125:300, 0.0625:300, 0.03:300 ng M:F) Mixture Ratio (male:female)
- 405 samples x 8 labs Single Source (Concordance)

---

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
**Summary of Literature Examined**

Reported Developmental Validation Efforts

<table>
<thead>
<tr>
<th>Kit</th>
<th>Reference</th>
<th>Terminology</th>
<th>Baseline</th>
<th>Mixture</th>
<th>Stutter</th>
<th>Precision</th>
<th>Sensitivity</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yfiler</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A total of 64 papers examined</td>
</tr>
<tr>
<td>Y-PLEX 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y-PLEX 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PowerPlex Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yfiler</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP ES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP 16 BIO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yfiler</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP 1.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SGM Plus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identifiler</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cofiler</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Profiler Plus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Full list of forensic DNA literature reviewed is available on STRBase

**Laboratory Internal Validation Summaries**

Soliciting Information on Studies Performed by the Community

http://www.cstl.nist.gov/biotech/strbase/validation.htm

**Acknowledgments**

- **NIJ Funding** for NIST Project Team through NIST Office of Law Enforcement Standards
- Co-Authors on Validation Work: Chris Tomsey and Margaret Kline
- Dave Duewer (NIST)
- Kari Tontarski (Montgomery County Crime Lab)
- Robin Cotton (Orchid Cellmark)
- Tim McMahon (AFDIL)
- Many members of forensic DNA typing community for their valuable input on our validation questionnaire

**Interlaboratory Studies**

DNA Quantitation (2004), Mixture Interpretation (2005)

**NIST Initiated Interlaboratory Studies**

<table>
<thead>
<tr>
<th>Studies involving STRs</th>
<th># Labs</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixture Interpretation Study (Jan-Mar 2005)</td>
<td>64</td>
<td>Data analysis currently on-going ... Will be presented at NIST Quantities and SQGDM (June 2005) and ISFG (Sept 2005)</td>
</tr>
</tbody>
</table>

**Individual Performance in an Interlaboratory Study**

DNA Quantitation Results from each laboratory are returned to them in comparison to other participating labs to illustrate opportunities for improvement...

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
NIST Quantitation Study 2004 (QS04)

Consisted of:
• 8 DNA extracts labeled A – H
• Shipped Dec 2003 – Jan 2004 to 84 laboratories for quantification; data received back by April 2004
• Labs were requested to use multiple methods / multiple analysts

We received data from 80 Labs (95%)

Total of 287 sets of data
Participants used 19 different quantification methods (primarily variations on Quantiblot and Real-time PCR)

Information from this interlab study is being used to help construct SRM 2372 (Human DNA Quantitation Standard)

8 DNA Samples in This NIST Study

Labs only being asked to provide their quant values (no typing results expected)

Volume of each DNA sample provided = 100 µL

1. Interlaboratory Comparisons

Laboratories are only being asked to provide their quant values (no typing results expected)

Mixed source DNA

Single source DNA

Teflon tube

Real-time qPCR Work at NIST

• Careful examination of published assays on the same set of DNA samples

• Lot-to-lot variability with Quantifiler "standard"
  – qPCR is a relative measurement that depends on the quality of the material used to generate the standard curve

84 laboratories were sent samples (80 returned results)

Participation in NIST Interlaboratory Study Participation in NIST Interlaboratory Study on DNA Quantitation (QS04)

AFDIL
FBI

Outside U.S.:
Germany – RCMP & CFS
South Africa
UK – FSS

Companies:
Applied Biosystems
Promega
Identity Genetics
Orchid Cellmark
BBI Biotech
Bode

Non-forensic labs:
NIH/NCI
ATCC

37 states + Puerto Rico
Variability of Quantifiler DNA Standards

Two lots of ABI “standards” using Quantifiler Human assay

<table>
<thead>
<tr>
<th>Sample (n = 4)</th>
<th>Standard Lot 1 (ng/mL)</th>
<th>Standard Lot 2 (ng/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4*</td>
<td>2.91 ± 0.04</td>
</tr>
<tr>
<td>2</td>
<td>7.26 ± 0.79</td>
<td>4*</td>
</tr>
<tr>
<td>3</td>
<td>2.93 ± 0.27</td>
<td>1.88 ± 0.09</td>
</tr>
<tr>
<td>4</td>
<td>3.46 ± 0.30</td>
<td>2.22 ± 0.08</td>
</tr>
<tr>
<td>5</td>
<td>2.99 ± 0.28</td>
<td>1.91 ± 0.08</td>
</tr>
<tr>
<td>6</td>
<td>2.62 ± 0.22</td>
<td>1.70 ± 0.03</td>
</tr>
</tbody>
</table>

* - indicates “standard” value based on starting material provided by the manufacturer
Samples 1, 3 = commercially available kit standards
Samples 4-6 = in-house standards based on UV absorbance

Mixture Interpretation Interlab Study (MIX05)

- Only involves interpretation of data
- 91 labs enrolled for participation (20 from overseas)
- 64 labs have returned results
- Four mock cases supplied with “victim” and “evidence” electropherograms (GeneScan .fsa files - that can be converted for Mac or GeneMapper; gel files made available to FMBIO labs)
- Data available with Profiler Plus, COIfiler, SGM Plus, PowerPlex 16, Identifier, PowerPlex 16 BIO (FMBIO) kits
- Summary of results will involve training materials to illustrate various approaches to solving mixtures