NIJ National Forensic Science Technology Center
President's DNA Initiative - Workshops

Validation Workshop

The Challenge of Inconsistency Between Laboratories in Validation

John M. Butler, PhD
National Institute of Standards and Technology (NIST)

Presentation Outline

Introductions: Presenters and Participants

Day #1
- Validation Overview (John)
- Introduction to DAB Standards (Robyn & John)
- Developmental Validation (John)

Day #2
- Inconsistency in Validation between Labs (John)
- Internal Validation (Robyn)
- Method Modifications and Performance Checks (Robyn)

Day #3
- Practical Exercises (Robyn)

Questions Asked in Summer 2004

- How consistent are various forensic laboratories in performing internal validation?
- Can validation be standardized and therefore made easier for forensic laboratories?

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

Contacting the Community

- **Validation Standardization Questionnaire** handed out at NIJ DNA Grantees meeting (June 28-30, 2004)
- Emails sent to >200 scientists (July-Aug 2004)
  - Attendees from the NIJ 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

Validation Standardization Questionnaire
Please return to John Butler (NIST): john.butler@nist.gov or 301-975-8505 (fax)

Purpose of questionnaire: We are embarking on an effort to define the minimum number of samples needed to reliably validate DNA typing procedures. As part of this effort, we are conducting a survey of standard practices currently used by practitioners in forensic DNA laboratories. Your honest responses to the following questions will help the entire community as we compile this information. Results will be summarized at the Promega meeting in October 2004 and made available on the NIST STRBase web site.

General Questions

What does the term validation mean to you? *(define in a single sentence if possible)*

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?
- ☐ Precision studies ___(indicate types of samples –i.e., ladders), # samples/run ____; # runs ___
- ☐ Sensitivity studies_____, what range? _____
- ☐ Mixture studies _____ what mixture ratios are needed? _____
- ☐ Non-human DNA studies _____
- ☐ Non-probative cases _____

How many total samples do you think it takes to internally "validate" a new forensic kit?
- ☐ 10
- ☐ 50
- ☐ 500
- ☐ Other: _____

Validation Standardization Questionnaire (conducted June-August 2004)

52 Survey Respondents

Individual	Lab Location	Individual	Lab Location	Individual	Lab Location
Abirami Chidambaram	AK	Janel Smith	CO	Martin Buoncristiani	CA
Ann Marie Gross	MN	Janice Nicklas	VT	Meghan Clement	NC
Bridget Tincher	WV	Jeff Ban	VA	Michael Hass	FL
Bruce McCord	FL	Joanne B. Sgueglia	MA	Neils Morling	Denmark
Carl Sobieralski	IN	Joe Mathew	TX	Paul Bush	IA
Carmen Tirado	PR	John Hartman	CA	Peg Scheartz	VT
Cary Maloney	MO	John P. Simich	NY	Sindey Schueler	KS
Cathryn Braunstein	MD	Joseph Abraham	CT	Steve LaBonne	OH
Cecilia A. Crouse	FL	Julia Naylor	LA	Terry Coons	OR
Charles Barna	MI	Julie Kempton	MD	Tim Kupferschmid	Myriad
David Einum	Orchid Cellmark	Ken Konzak	CA	Tom Scholl	Myriad
Earl Ritzline	FL	Kris Radecki	NM		
Eric Buel	VT	Kris Whitman	AZ		
Farida Alshamali	Dubai	Larry Blanton	CA		
Gary Shutler	WA	Linda Jankowski	NJ		
George Schiro	LA	Lisa Dowler	MO		
Hope Olson	ND	Marcia LaFountain	VT		
James Schumm	Bode	Mark Squibb	OH		

5 anonymous individuals

Responding after Promega meeting
George Duncan (FL)
Joseph Galdi (NY)

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 Identifiler with ABI 310 and ABI 3100; extensive court experience
- **AFDIL** – 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…

Validation Standardization Questionnaire (conducted June-August 2004)

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 …

How I felt after taking on this project…

Me

Literature, Validation Data, Survey Responses

Validation Standardization Questionnaire (conducted June-August 2004)

How do you know when you are finished with a validation study? (1)

- "When you have demonstrated that it works as expected over a range of samples that is representative of what is seen in casework"
- "When repeat performance gave the same result"
- "**When you pull the toothpick out and it is dry?...** Meet at least minimum expectations and DAB guidelines"
- "You are very comfortable that you know how it works and your documentation will convince a reviewer you have put the kit thru a rigorous review/test."

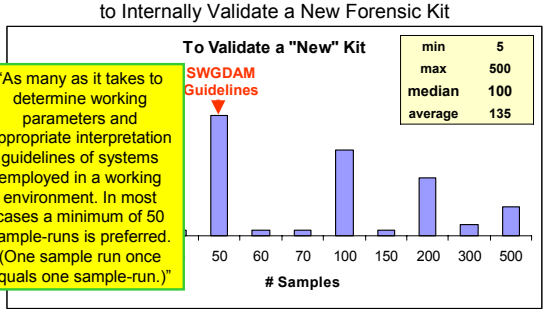
Validation Standardization Questionnaire (conducted June-August 2004)

How do you know when you are finished with a validation study? (2)

- "Once a reasonable body of data has been assembled and analyzed, quirks have been revealed, and the upper and lower limits of the system have been challenged using a range of samples that one could expect to encounter in the everyday operation of the system"
- "When you achieve accuracy and precision to the desired statistical level of certainty"
- "You can never know…but it is always nice to have more samples!"
- "Validation is never complete"

Validation Standardization Questionnaire (conducted June-August 2004)

Survey Summary for Recommended **Total Number of Samples** to Internally Validate a New Forensic Kit



"As many as it takes to determine working parameters and appropriate interpretation guidelines of systems employed in a working environment. In most cases a minimum of 50 sample-runs is preferred. (One sample run once equals one sample-run.)"

Choices in survey were: **10, 50, 500, or other ____**

Validation Standardization Questionnaire (conducted June-August 2004)

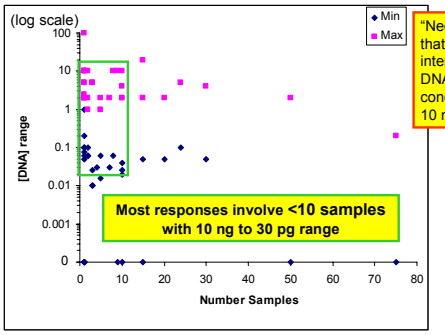
Survey Summary for Recommended **Precision Studies**

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)

Validation Standardization Questionnaire (conducted June-August 2004)

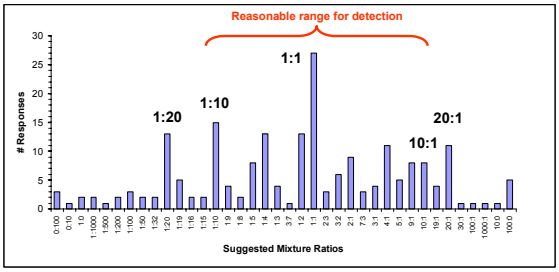
Survey Summary for Recommended **Sensitivity Studies**



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

Validation Standardization Questionnaire (conducted June-August 2004)

Survey Summary for Recommended **Mixture Studies**



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

Validation Standardization Questionnaire (conducted June-August 2004)

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."

Validation Standardization Questionnaire (conducted June-August 2004)

Survey Summary for Recommended **Non-Probative Cases**

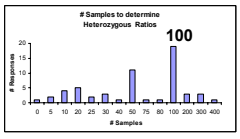
A few of the responses:

- 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."

Validation Standardization Questionnaire (conducted June-August 2004)

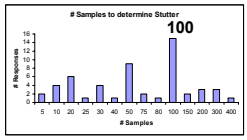
Survey Summary for Recommended **Numbers of Samples**

to Determine Heterozygote Peak Height Ratios and Stutter Values



min	0
max	400
median	50
average	85

Heterozygote Peak Height Ratios



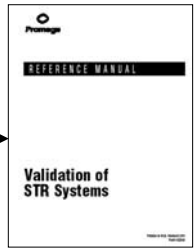
min	5
max	400
median	63
average	88

Stutter Values

Validation Standardization Questionnaire (conducted June-August 2004)

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



Published in March 2001

Validation Standardization Questionnaire (conducted June-August 2004)

Can Validation be Standardized?

Statements from survey responders…

Over 86% (45/52) said yes

Those who responded "no" said

- "to some degree it can be, however, validation is specific to the platform, kits, …",
- "a start-up lab should do much more than an experienced lab…",
- "validation builds on previous work by lab or published data",
- "parts of it can be standardized; I don't think the non-probative cases could be", and
- "only in a general way, as with the SWGDAM guidelines. The uniqueness of each new procedure would make standardization difficult."

Our Conclusion…

to a certain extent it can…but everyone will always have a different comfort level…and **inflexible, absolute numbers for defined studies will not likely be widely accepted**

Validation Standardization Questionnaire (conducted June-August 2004)

If a Standard Protocol or Set of Guidelines Existed for Validation, Would You Use It?

90% (47/52) said yes

Some responses

- "No-I would reference them. I may not completely abide by them but I would certainly review them",
- "No-but it would be taken into consideration",
- "Yes-we would have to or there would be problems in court",
- "Yes-as long as they remain updated, relevant and feasible guidelines and do not become dogma",
- "Yes-if it would pass an audit for validation", and
- "Yes-unless they were far less stringent than current practice."

Validation Standardization Questionnaire (conducted June-August 2004)

If a Standard Set of Samples Existed for Performing Validation Testing, Would You Use Them?

90% (47/52) said yes

Some responses

- "Yes-would love to have something like that available; we are always eager to have benchmarks for assessment",
- "Yes-these types of samples would cut down on time for validation. It would be efficient if they were ready for the particular type of validation…",
- "Yes-as long as they are readily available at a reasonable price",
- "No-this approach is not recommended. It is most important that systems work with the materials available in individual laboratories. Laboratories should be allowed, even encouraged, to select their own preferred materials. Choices for such selection of standard materials for within laboratory analyses and cross-laboratory comparison already exist from a variety of government and commercial entities."

There are Different Opinions… in Who Should Perform Validation

Development of New STRs for Forensic Casework: Criteria for Selection, Sequencing & Population Data and Forensic Validation

Angel Carracedo and M.V. Lareu
Institute of Legal Medicine. University of Santiago de Compostela, Spain

http://www.promega.com/geneticidproc/ussymp9proc/content/21.pdf

Validation studies following similar parameters to those recommended by TWGDAM were carried out. These include robustness, stability, mixtures, non-human studies, mutation rate and checking for independence with other loci. In our opinion the final validation of a system cannot be carried out by individual groups and companies and should always be performed by an internationally established validation group. In Europe a final assessment and intercomparison exercises are usually performed by the EDNAP group, a working group of the ISFH.

Abstract from talk presented at Promega meeting in 1998

Validation Section of the DNA Advisory Board Standards
issued July 1998 (and April 1999); published in *Forensic Sci. Comm.* July 2000

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**.

FORENSIC SCIENCE COMMUNICATIONS JULY 2000 VOLUME 2 NUMBER 3

Revised SWGDAM Validation Guidelines
(July 2004)

http://www.fbi.gov/hq/lab/fsc/backissu/july2004/standards/2004_03_standards02.htm

Forensic Science Communications July 2004 – Volume 6 – Number 3
Standards and Guidelines

Revised Validation Guidelines

Table of Contents
Back Issues
Search
Editors
About FS
Instructio
Authors

Scientific Working Group on DNA Analysis Methods (SWGDAM)

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

Program for DNA Analysis by the Technical Working Group on DNA Analysis Methods (*Crime Laboratory Digest* 1995;22(2):21-43) has been revised due to increased laboratory experience, the advent of new technologies, and the issuance of the Quality Assurance Standards for Forensic DNA Testing Laboratories by the Director of the FBI (*Forensic Science Communications* available: www.fbi.gov/hq/lab/fsc/backissu/july2000/codis2a.htm).

The document provides validation guidelines and definitions approved by SWGDAM July 10, 2003.

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."

Effort

Lots of experiments are required

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

Time

Validation Standardization Questionnaire (conducted June-August 2004)

Survey Summary of Planned Near-term "Validation"

Commercial Kits

Extraction
- DNA IQ
- Qiagen
- Biomek 2000

DNA Quant
- **Quantifiler**

STR Amp Kits
- Identifiler
- PowerPlex Y
- Yfiler
- PowerPlex 16
- ProPlus/COfiler reduced volume

Software
- **GeneMapper*ID***
- GeneScan/ Genotyper NT
- TrueAllele
- SQL*LIMS and Forensic Solution

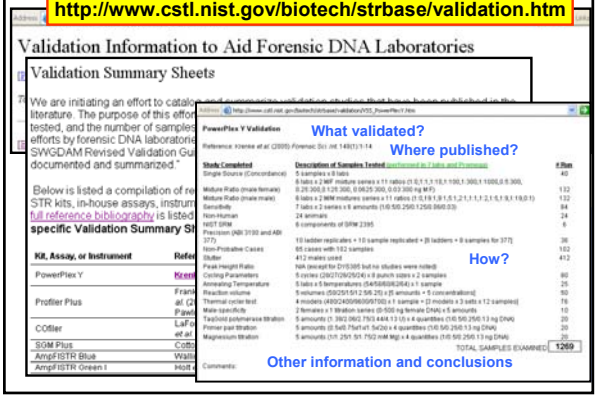
Analysis Instruments
- **ABI 3100 Avant**
- ABI 3100
- FMBIO III+
- MegaBACE

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

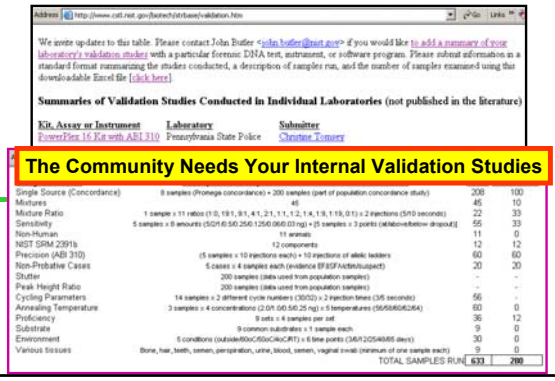
The ones in bold were most common

New Validation Homepage on STRBase

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



Laboratory Internal Validation Summaries



We invite updates to this table. Please contact John Butler <john.butler@nist.gov> if you would like to add a summary of your laboratory's validation studies with a particular forensic DNA test, instrument, or software program. Please submit information in a standard format summarizing the studies conducted, a description of samples run, and the number of samples examined using this downloadable Excel file [click here].

Summaries of Validation Studies Conducted in Individual Laboratories (not published in the literature)

Kit, Assay or Instrument	Laboratory	Submitter
PowerPlex 16 Kit with ABI 310	Pennsylvania State Police	Christine Tomsey

The Community Needs Your Internal Validation Studies

Single Source (Concordance)	8 samples (Promega concordance) + 200 samples (part of population concordance study)	208	100
Mixtures	45	45	10
Mixture Ratio	1 sample x 11 ratios (1:0, 19:1, 9:1, 4:1, 2:1, 1:1, 1:2, 1:4, 1:9, 1:19, 0:1) x 2 injections (5/10 seconds)	22	33
Sensitivity	5 samples x 8 amounts (SQ/1:0.5/0.25/0.125/0.06/0.03 ng) + [5 samples x 3 points (at/above/below dropout)]	55	33
Non-Human	11 animals	11	0
NIST SRM 2391b	12 components	12	12
Precision (ABI 310)	(5 samples x 10 injections each) + 10 injections of allelic ladders	60	60
Non-Probative Cases	5 cases x 4 samples each (evidence [FBI/SFA/other/suspect]	20	20
Stutter	200 samples (data used from population samples)	-	-
Peak Height Ratio	200 samples (data used from population samples)	-	-
Cycling Parameters	14 samples x 2 different cycle numbers (30/32) x 2 injection times (3/5 seconds)	56	-
Annealing Temperature	3 samples x 4 concentrations (2.0/1.0/0.5/0.25 ng) x 5 temperatures (56/58/60/62/64)	60	0
Proficiency	9 sets x 4 samples per set	36	12
Substrate	9 common substrates x 1 sample each	9	0
Environment	5 conditions (outside/60oC/50oC/4oC/RT) x 6 time points (3/6/12/OS/4W/6 days)	30	0
Various tissues	(Bone, hair, teeth, semen, perspiration, urine, blood, semen, vaginal swab (minimum of one sample each)	9	0
	TOTAL SAMPLES RUN	633	280

Acknowledgments

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

- **NIJ Funding** for NIST Project Team through NIST Office of Law Enforcement Standards
- **Co-Authors: 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**