

Preclinical validation of fluorescence in situ hybridization assays for clinical practice

Anne E. Wiktor, BS, Daniel L. Van Dyke, PhD, Peggy J. Stupca, MS, Rhett P. Ketterling, MD, Erik C. Thorland, PhD, Brandon M. Shearer, BS, Stephanie R. Fink, BA, Kimberly J. Stockero, BS, Jason R. Majorowicz, BS, and Gordon W. Dewald, PhD

Purpose: Validation of fluorescence in situ hybridization assays is required before using them in clinical practice. Yet, there are few published examples that describe the validation process, leading to inconsistent and sometimes inadequate validation practices. The purpose of this article is to describe a broadly applicable preclinical validation process. **Methods:** Validation is performed using four consecutive experiments. The Familiarization experiment tests probe performance on metaphase cells to measure analytic sensitivity and specificity for normal blood specimens. The Pilot Study tests a variety of normal and abnormal specimens, using the intended tissue type, to set a preliminary normal cutoff and establish the analytic sensitivity. The Clinical Evaluation experiment tests these parameters in a series of normal and abnormal specimens to simulate clinical practice, establish the normal cutoff and abnormal reference ranges, and finalize the standard operating procedure. The Precision experiment measures the reproducibility of the new assay over 10 consecutive working days. To illustrate documentation and analysis of data with this process, the results for a new assay to detect fusion of *IGH* and *BCL3* associated with t(14;19)(q32;q13.3) in lymphoproliferative disorders are provided in this report. **Results:** These four experiments determine the analytic sensitivity and specificity, normal values, precision, and reportable reference ranges for validation of the new test. **Conclusion:** This report describes a method for preclinical validation of fluorescence in situ hybridization studies of metaphase cells and interphase nuclei using commercial or home brew probes. **Genet Med 2006;8(1):16–23.**

Key Words: FISH validation, metaphase FISH, interphase FISH

Validation of fluorescence in situ hybridization (FISH) assays is becoming more challenging as the number of probes and applications increase, and diverse analytic strategies emerge. The Clinical Laboratory Improvement Amendments (CLIA), Food and Drug Administration (FDA), College of American Pathologists, and other accrediting agencies all require validation of new or modified FISH assays before reporting any patient results. Preclinical validation requires evaluation of the accuracy, analytic sensitivity and analytic specificity (interfering factors), normal values, precision, and reportable reference ranges of the FISH assay.^{1–3} This publication focuses on the preclinical validation process, but validation continues into clinical practice and must be continually monitored to ensure the FISH assay works as expected and achieves the intended results. In clinical practice, validation includes proficiency testing, assessment of employee competency, instrument calibration, and correlation with clinical findings.

Some regulatory agencies, such as the College of American Pathologists and the New York State Health Department, provide general standards for validation of FISH tests.^{2,4} The American College of Medical Genetics (ACMG) has published guidelines to establish scoring criteria, analytic sensitivity, analytic specificity, normal cutoffs, and abnormal reference ranges.⁵ The ACMG has also published a policy statement discussing the clinical considerations of FISH for prenatal screening, diagnosis of microduplication and microdeletion syndromes, and identification of acquired marker or derivative chromosomes.⁶

In 1995, Schad and Dewald⁷ presented an overview of quality control and quality assurance methods, and provided suggestions on validation methods for new FISH tests. Subsequently, Cohen et al.⁸ described standardization criteria including quality control and quality assurance methods for detection of *BCR/ABL* fusion in interphase nuclei. In 2003, Hausmann and Cremer⁹ led a workshop that dealt with standardization of FISH procedures for clinical practice. Other authors have also emphasized the need for test validation and quality control and quality assurance monitoring.^{10,11} In a series of three independent studies, a group of cytogenetic laboratories worked together to evaluate the efficacy of FISH proficiency testing on metaphase and interphase

From the Division of Laboratory Genetics, Mayo Clinic, Rochester, Minnesota.

Gordon W. Dewald, Division of Laboratory Genetics, Mayo Clinic, 200 First St SW, Rochester, MN 55905.

Received: July 26, 2005.

Accepted: October 31, 2005.

DOI: 10.1097/01.gim.0000195645.00446.61

preparations, and identified methods to calculate specific analytic parameters.¹²⁻¹⁴ Dewald¹⁵ published a book chapter on a procedure for validation of *BCR/ABL* fusion in interphase nuclei; this method has considerable application for other quantitative interphase FISH assays as well. In 2004, the National Committee on Clinical Laboratory Standards (NCCLS) (now known as the Clinical and Laboratory Standards Institute) published a comprehensive description of processes to validate FISH assays.¹⁶

Despite all published requirements and guidelines for validation of FISH assays, step-by-step details to validate new FISH assays are rare in the literature. Consequently, validation of FISH assays is inconsistent among laboratories and inadequate in some. Here, we describe a systematic procedure that involves four experiments (familiarization, pilot study, clinical evaluation, and evaluation of precision) to validate FISH assays based on the NCCLS guidelines.¹⁶ This procedure focuses on preclinical processes and is applicable to conventional FISH studies of either metaphase cells or interphase nuclei. To illustrate documentation and analysis of data with this process, the results for a new dual-color/double-fusion (D-FISH) assay to detect fusion of *IGH* and *BCL3* associated with *t(14;19)(q32;q13.3)* in lymphoproliferative disorders are provided in this report.

METHODS

Experiment 1: Familiarization

Purpose

Experiment 1 gains initial experience with performance of the FISH test and determines the analytic specificity and sensitivity of the assay for peripheral blood samples.

Experiment

The FISH probe is hybridized to metaphase and interphase cells from peripheral blood cultures of five karyotypically normal males. For each specimen, a technologist evaluates the target chromosome loci from 20 consecutive intact metaphases and records the signal patterns of 50 consecutive interphase nuclei. The number of FISH signals in each cell is recorded, and the hybridization sites in metaphase cells are identified by chromosome morphology (chromosome size, centromere index, and reverse banding using either DAPI or sequential G-banding to FISH). An overall impression of probe performance and equipment is documented, and images of representative cells are collected. The analytic sensitivity and specificity for metaphase cells, and the percentage of nuclei that meet the signal pattern criteria for normal cells are calculated. The results are interpreted and summarized as shown in Table 1.

Experiment 2: Pilot study

Purpose

Experiment 2 gains experience with the FISH assay for normal and abnormal specimens using the tissue for which the test

is intended and establishes preliminary scoring criteria, normal cutoff, analytic sensitivity, and performance of the FISH assay.

Experiment

Five normal and five representative abnormal specimens from the intended tissue type (including variant abnormalities) are analyzed. Previously validated FISH assays that use similar probe strategies are reviewed to establish initial scoring criteria and the number of cells to analyze. The scoring criteria should include all expected normal and abnormal patterns. These criteria are used to score consecutive qualifying interphase nuclei (two technologists evenly divide the effort of scoring) and/or metaphase cells (one technologist) from each specimen. The signal pattern of each cell is recorded, and images of representative cells are documented. The signal patterns observed are compared with similar expected previously validated FISH strategies. The normal cutoff and analytic sensitivity for normal specimens are calculated, and the percentage of cells that meet the scoring criteria is established. A preliminary standard operating procedure is written. All results are interpreted and summarized as shown in Table 2.

Experiment 3: Clinical evaluation

Purpose

Experiment 3 tests predictable parameters encountered in clinical practice, establishes the normal cutoff and abnormal reference range, and validates the standard operating procedure.

Experiment

Test 25 normal samples and a series of abnormal specimens to simulate clinical practice. If applicable, include samples with variants of the chromosome anomaly and samples with various proportions of normal and abnormal cells to test detection of residual disease. Code and randomize specimens. Two technologists (or one technologist for metaphases) should independently score each specimen using scoring criteria from the pilot study. For each specimen, cells with signal patterns that do not meet the scoring criteria are recorded and investigated because these may identify a new scoring pattern. Images of representative cells from abnormal specimens are documented. At the end of the study, the samples are unblinded and the results are correlated with the diagnostic "gold standard" (e.g., karyotype, flow cytometry, molecular genetic studies, patient phenotype, or disease status). The results are interpreted and summarized as shown in Table 3.

Experiment 4: Precision

Purpose

Experiment 4 tests the reproducibility (precision) of the FISH assay.

Table 1
Familiarization: 5 normal male peripheral blood samples

Probe assay								
<i>BCL3</i> (19q13.3) SpectrumOrange™ (home brew) – 570 kb								
<i>IGH</i> (14q32) SpectrumGreen™ (Vysis, Inc.) – 1.5 Mb								
Lot No.: 123456				Date: 7-2-2004			Technologist: KS	
Assay results								
Pt	Interphase nuclei ^a			Unexpected patterns	Metaphase cells			
	Expected patterns ^b				Orange <i>BCL3</i>		Green <i>IGH</i>	
	2R2G	1R1G1F	Other		19q13.3	Other	14q32	Other
1	40	5	4	1 (5R2G)	20		20	
2	41	4	5	0	20		20	
3	39	6	4	1 (1R2F)	19	1 ^c	20	
4	42	5	3	0	20		20	
5	41	5	4	0	20		20	
Interpretation								
Statistic	Interphase				Metaphase			
Analytical specificity	na ^d				<i>BCL3</i> 99%; <i>IGH</i> 100%			
Analytical sensitivity	na				<i>BCL3</i> 100%; <i>IGH</i> 100%			
% of all cells meeting scoring criteria	99%				na			

^aFor clarity in recording signal patterns, Spectrum Green™ will be designated with a “G”, Spectrum Orange™ will be designated with an “R”, and a fusion of green and orange will be designated with an “F.” For example: one green signal, one orange signal and two fusion signals will be described as “1G1R2F.”

^bIncludes patterns seen due to signal overlap and random loss (i.e. 1R1G1F, 1R2G, 2R1G, 1R1G, 2F).

^cOne metaphase had correct localization at 19q13.3 on both chromosome 19s and cross-hybridization of *BCL3* to chromosome 2p.

^dna, not applicable.

Experiment

Select a specimen with a known proportion of normal and abnormal cells. Perform and analyze FISH studies on this specimen on 10 consecutive days. The precision is calculated as the mean, standard deviation, and range of the results over 10 days. These statistics represent the potential of the FISH assay to accurately determine the percentage of abnormal cells.

RESULTS

Experiment 1: Familiarization

Study phytohemagglutinin-stimulated metaphases

In clinical practice, FISH tests are applied to different tissues and cell types, but it is useful to initially establish the analytic specificity and sensitivity for metaphase cells and to evaluate interphase nuclei from phytohemagglutinin-stimulated peripheral blood cultures from normal males. This experiment permits direct comparison of assay performance for different FISH probes regardless of the final application of the FISH assay.

Selection of cells to be scored and interpreting signal patterns

The criteria for selection of cells to score should be established. In general, metaphase cells are scorable if they appear intact and chromosomes are sufficiently free of overlap to confidently confirm location of the hybridization signal. For interphase nuclei, it is best to score nonoverlapping nuclei.¹⁷

Guidelines to define fusion, break-apart, and overlapping signals should be established to ensure consistent analyses.¹⁷ This information is often available from other validated FISH assays, but to ensure the new FISH test performs correctly, the technologists should record all signal patterns in a consecutive series of cells.¹⁸

Analytic sensitivity

Analytic sensitivity of the test can be defined as the percentage of chromosome targets or interphase nuclei with the expected signal pattern. For example, in analysis of normal metaphase analysis, if 99 of 100 chromosome targets show the expected normal signal pattern then the analytic sensitivity is 99%. Determination of analytic sensitivity for interphase nuclei is less accurate than for metaphase because of uncertainty of where signals hybridize, overlapping signals, signal integrity,

Table 2
Pilot Study—Interphase testing (Data from 4 of the 10 study samples are shown)

Probe assay								
Probe: <i>IGH/BCL3</i>			Lot#: 123456			Date: 7-6-2004		
Strategy: D-FISH fusion of <i>IGH/BCL3</i> at 14q32 and <i>BCL3/IGH</i> at 19q13.3								
Assay results								
Pt	Tech	Cells meeting scoring criteria					Cells not meeting scoring criteria	Total cells analyzed
		2R2G	1R1G1F	1R1G2F	Other	Total		
1	KS	235	15	0	0	250	25	275
	SP	232	17	0	1 (2R3G)	250	40	290
2	KS	240	10	0	0	250	31	281
	SP	233	17	0	0	250	50	300
9	KS	55	15	0	180 (2R3G)	250	23	273
	SP	70	13	0	167 (2R3G)	250	28	278
10	KS	35	5	210	0	250	23	273
	SP	35	7	208	0	250	13	263

Interpretation:
 1. Scoring criteria: 2R2G, 1R1G1F, 1R1G2F, 2R2G1F, 1R2G1F, 2R1G1F, 2R3G
 2. Maximum false-positive nuclei for normal specimens = 1
 3. Initial cutoff for normal specimens: e.g. 1R1G2F = 0.6%; 2R3G = 1.0%
 4. Percent of all cells meeting scoring criteria: 89.7%
 5. Overall analytical sensitivity for normal specimens: 98.5%

Table 3
Clinical Evaluation Summary

Probe assay							
Probe: <i>IGH/BCL3</i>		Lot #: 123456		Date: 7-20-2004			
Assay results							
Number of cells scored/specimen				500			
% of all cells meeting scoring criteria				89%			
Scoring criteria modified?				Yes			
Unexpected patterns?				Yes (see below)			
No. of normal specimens				25			
% normal specimens identified as true normal				100%			
No. of abnormal specimens				24			
% abnormal specimens identified as true abnormal				100%			
Interpretation							
Statistic		2R3G	1R1G1F	all other patterns			
Max. false-positive cells from normal specimens		8	35	0			
Normal cutoff (%)		2.83	8.58	0.60			
.....							
Signal pattern		2R3G	1R1G2F	2R1G1F	2R1G2F	1R2G2F	1R1G3F
Reference range (%) for abnormal specimens		37–64	12–93	83	98	72–92	56–84

and other technical factors. Therefore, we recommend calculating the percentage of interphase nuclei that meet the scoring criteria. We calculate analytic sensitivity as the percentage of

interphase nuclei that have the expected normal signal pattern among cells that meet any of the expected signal patterns in normal subjects.

Table 4
Normal cutoff calculated with the binomial expansion formula^a

False positive cells ^c	Number of cells scored ^b				
	50	100	200	500	6000
0	6	3	1.5	0.60	0.050
1	10	5	2.5	0.95	0.079
2	12	7	3.5	1.30	0.105
3	14	8	4.0	1.54	0.129
4	18	9	4.5	1.81	0.152
5	20	10	5.5	2.07	0.175
6	22	11	6.0	2.33	0.197
7	22	12	6.5	2.58	0.219
8	24	13	7.0	2.83	0.240
9	26	14	7.5	3.07	0.261

^aThe normal cutoff values in this table have been rounded off to match the precision that will be achieved for the number of cells scored. For example, scoring 50 cells permits precision of only 1/50 (2% increments) whereas analysis of 6000 cells permits precision of 1/1000 (0.001% increments).

^bThese calculations assume that cells have been randomly selected for analysis; this is best done by scoring a consecutive series of cells that meet the scoring criteria. ^cNormal cutoff: In our experience, D-FISH strategies usually have zero false-positive cells for the classical signal pattern (1R1G2F), resulting in a normal cutoff of >0.6% for 500 nuclei and <0.05% for 6000 nuclei. For comparison, breakapart FISH (BAP-FISH) strategies usually have 4 false-positive cells for the classical signal pattern (1R1G1F), resulting in a normal cutoff of 4.5% for 200 nuclei. Extra-signal FISH (ES-FISH) strategies usually have 4 false-positive cells for the classical signal pattern (2R1G1F) for 200 nuclei, resulting in a normal cutoff of 4.5%. Enumeration FISH strategies have 4 to 9 false-positive cells among 200 nuclei for the classical deletion signal pattern (1R2G), resulting in a normal cutoff of 4.5 to 7.5%.

Hybridization signals on metaphase cells can be analyzed at either the chromosome or chromatid level. In our experience, scoring signal patterns at the chromosome level is more reliable because of fewer interfering factors. Because chromatids often twist and overlap, scoring at the chromatid level may interfere with assessment of the analytic sensitivity.

Analytic specificity

Analytic specificity is defined as the percentage of FISH signals at the expected target locus and no other chromosomal location. For example, if 99 FISH signals are observed at the expected chromosome location and 1 FISH signal is seen at an incorrect chromosomal location, then the analytic specificity is 99%. The analytic specificity for interphase nuclei is not calculated because it is not readily apparent whether the FISH signals have hybridized to the expected target locus. Most clinical FISH assays have an analytic specificity greater than 98%; if probe performance is less than optimal then modification of the probe may be necessary.

Some FISH probes cross-hybridize with other loci. For example, commercial chromosome 15 centromere probes cross-hybridize with the centromere region of chromosome 14 in some individuals. It is important to determine whether any new FISH probe cross-hybridizes with unexpected loci. To determine whether a probe cross-hybridizes with the Y chromosome, it is necessary to use normal male specimens in this experiment.

Background signals differ from cross-hybridization because they are nonspecific and randomly distributed on the microscope slide. Background signals can be associated with poor sample and slide preparation, poor hybridization, incorrect

washing stringencies, or other factors. Background signals may be acceptable for some metaphase studies because it is possible to visualize the chromosome targets, but can be problematic for analysis of interphase nuclei because the hybridization locus is uncertain.

Equipment assessment

The equipment (e.g., water baths, micropipettes, and incubators) used to hybridize FISH probes with genomic DNA must be adequate to perform the test consistently and reproducibly. Microscopes, filters, and imaging systems are assessed to determine whether they are adequate for analysis and suitable for documentation of results in clinical practice.

Illustration of familiarization results for IGH and BCL3

Table 1 provides results of a familiarization experiment done on blood from five normal males who were studied with a new FISH assay using DNA probes for *BCL3* and *IGH*. The probe name, chromosomal hybridization location, size of probe, and manufacturer of the probe are documented. The technologist, date, and lot number of FISH probes studied are provided. The results obtained for each patient are documented for interphase and metaphase cells. The number of interphase nuclei with the expected or unexpected signal patterns is listed. The number of metaphases in which the probe hybridized to the correct locus or incorrect locus is provided. In one metaphase, a *BCL3* signal hybridized at 19q13.3 on both chromosomes 19 and at an errant location at 2p. Thus, the analytic specificity for metaphase cells is 99% for *BCL3*. Because the *IGH* signal did not hybridize to any locus other than 14q32 in any metaphase, the analytic specificity for *IGH* is

100%. Because *BCL3* and *IGH* hybridized to their correct locus in each of the 100 metaphases, the analytic sensitivity is 100% for *BCL3* and *IGH*. Overall, 99 of the 100 metaphases met the expected scoring criteria for this FISH assay. Thus, the percent of interphase nuclei that met the scoring criteria is 99%.

Experiment 2: Pilot study

Tissue type

CLIA requires validation for each type of tissue that is intended to be tested in clinical practice.³ Similar results from different tissue types can be combined, but they should first be evaluated independently. For example, bone marrow and peripheral blood samples may be used to validate probes for hematologic malignancies, and the results may be combined if they are similar. Likewise, metaphase cells from amniotic fluid and chorionic villi may be used to validate FISH probes for 22q11.2 deletion, and the results may be combined if they are similar.

For some tissues, validation depends on the cell type to be analyzed. For example, cells processed with both fluorescent immunophenotyping markers and FISH probes require immunophenotype positive cells to be scored.¹⁹ Urologic samples require consideration of cell morphology.²⁰ Paraffin-embedded tissue sections should be cut to preserve whole nuclei and cells selected to minimize analysis of overlapping nuclei.

Interpreting signal patterns

The scoring criteria of metaphase and interphase cells in this experiment can be defined on the basis of the experience with other probe sets using the same strategy. In particular, note signal integrity, size, overlap, and other characteristics of signals. The frequency of each expected signal pattern and any new pattern should be documented and images collected for the record. For example, with D-FISH strategies in clinical practice, we routinely analyze 500 interphase nuclei to detect very low levels of disease. Thus, to validate a D-FISH probe, two technologists would score a consecutive series of cells until 500 nuclei meet the strict scoring criteria.

The results of this experiment may expose technical and clinical variations of the new FISH assay, and develop initial scoring criteria for clinically relevant signal patterns. It is important to examine the signal patterns of nuclei that did not meet the scoring criteria to determine whether they reflect technical problems or signal patterns from an unexpected clinically significant clone.

Number of cells to be analyzed

The number of cells to be analyzed should meet the needs of the new FISH assay in clinical practice. Some guidance on this issue can be gleaned from ACMG and NCCLS guidelines.^{5,16} Statistical models can be used to project the number of consecutive cells needed to achieve a certain analytic sensitivity with a particular degree of confidence. Table 3 shows the relationship between the number of false-positive cells and the analytic sen-

sitivity of the assay to detect a second cell population with 95% confidence.

Normal cutoff for each signal pattern

Preliminary normal cutoff values for each of the common signal patterns may be available from experience and should be confirmed by analysis of results from the pilot study. The normal cutoff is calculated by using the maximum number of false-positive cells for any normal sample and using a binomial statistical formula to project the upper bound of the 95th percentile.¹³ This computation will help predict whether the new FISH assay will meet clinical expectations (more details on determining the normal cutoff are provided in Table 4 and in the clinical evaluation study).

Analytic sensitivity

For interphase assays, analytic sensitivity can be calculated for normal specimens. As an illustration, if all the cells examined have a normal signal pattern and no cells have an abnormal signal pattern, then the analytic sensitivity is 100%. If the FISH assay is for metaphase cells, then the analytic sensitivity can be calculated for both normal and abnormal specimens.

Laboratory procedure and safety precautions

After completion of the pilot study, a provisional standard operating procedure is written and a safety check of reagents, space, and equipment is performed.¹⁶

Results for pilot study with *IGH* and *BCL3*

Table 2 illustrates results from a pilot study of *IGH* and *BCL3* FISH probes on interphase nuclei. This series included bone marrow or blood from five normal individuals, two patients with t(14;19)(q32;q13.3), and three patients with t(14;18)(q32;q21.3). Because this new assay used a D-FISH strategy to detect fusion of *IGH* and *BCL3*, in this experiment we initially applied our standard scoring criteria from other D-FISH methods because they often perform in a similar manner. Two technologists scored 250 nuclei for each sample to produce results for a total of 500 nuclei. This FISH probe set was designed to produce typical D-FISH signal patterns including 2R2G, 1R1G1F, 1R1G2F, 2R2G1F, 1R2G1F, and 2R1G1F. Typical signal patterns and scoring criteria for various FISH strategies have been decided by Dewald and colleagues.¹⁸ The results in Table 2 are shown for 4 of 10 subjects we evaluated in this experiment. The number of nuclei that met or did not meet the scoring criteria is provided. Only one "unexpected" signal pattern was encountered in this experiment, namely, 2R3G. This signal pattern was attributed to separation of the *IGH* signal in patients with t(14;18)(q32;q21.3) and added a new scoring criterion for this assay. In this experiment, the initial cutoff for 2R3G was 1.0% and 0.6% for all other signal patterns in the scoring criteria. The overall percentage of nuclei that met the scoring criteria for both normal and abnormal specimens was 89.7%. The overall analytic sensitivity for normal specimens was 98.5%.

Experiment 3: Clinical evaluation

Normal cutoff

The normal cutoff can be calculated in various ways,^{14,21} but because the results do not fit a Gaussian distribution it is incorrect to use the mean and standard deviation. A more appropriate statistical approach is the Microsoft Excel (Microsoft, Redmond, WA) beta inverse function, = BETAINV(confidence level, false-positive cells plus 1, number of cells analyzed), to calculate a one-sided upper confidence limit for a percentage proportion based on an exact computation for the binomial distribution (Table 4). To do this, examine the results for the first 20 normal specimens and identify the specimen with the greatest number of false-positive nuclei for any given signal pattern. This number can be used in the beta inverse function to determine the normal cutoff for detection of a true abnormal clone. To illustrate how to calculate the normal cutoff, consider the following example for a 95% confidence level in which four false-positive cells for any given signal pattern were identified among 500 nuclei. In the formula bar in Microsoft Excel, enter: = BETAINV(0.95,5,500); the result is 1.81% cutoff or 9.05 cells. In other words, the formula would read = BETAINV(0.95 upper bound percentile, 4 false-positive cells plus 1, analysis of 500 cells). On the basis of this calculation, the abnormal cutoff is 10 cells because fractions of cells cannot be analyzed. Thus, the observation of 10 or more cells in 500 cells analyzed would be an abnormal result. Once the normal cutoff has been established, results of the remaining specimens, including the five remaining normal specimens, can be used to test the validity of the cutoff for the new test.

Abnormal reference range

The abnormal reference range is defined as the lowest and highest percentage of cells with an abnormal signal pattern for patients with untreated disease. For a more informative reference range, whenever possible at least five abnormal samples should be analyzed. Depending on the study, the abnormal samples are from patients diagnosed with the congenital syndrome, with newly diagnosed neoplastic disorders, or with solid tumor malignancies. This information is useful in clinical practice to discriminate between specimens from patients with newly diagnosed disease and posttreatment specimens.

Correlation with “gold standard”

A comparison of FISH results with the diagnostic “gold standard” should be performed to determine the accuracy of the test (i.e., if all normal and abnormal samples were correctly identified with the new FISH assay). The percentages of abnormal cells may vary because of sampling error and other factors, but each specimen should be correctly identified as normal or abnormal. Any “gold standard” abnormal sample that produces FISH results that are not included in the strict scoring criteria should be investigated. Such specimens may define new scoring patterns.

Experimental clinical sensitivity and specificity

The experimental clinical sensitivity and specificity are estimates of the accuracy of the test. The experimental clinical sensitivity is defined as the percentage of correctly identified true-positive cases. The experimental clinical specificity is defined as the percentage of correctly identified true negative cases. The actual clinical sensitivity and specificity may require a much larger study group.

Standard operating procedure

At the completion of the clinical evaluation study, the provisional standard operating procedure should be amended to include the scoring criteria, cutoff values, and abnormal reference ranges. Any major procedural changes identified in the evaluation study may need to be further validated.

Results for clinical evaluation of IGH and BCL3

Table 3 summarizes the results for a clinical evaluation experiment based on 25 normal individuals, 6 patients with t(14;19)(q32;q13.3), 17 patients with t(14;18)(q32;q21.3), and 1 patient with t(11;14)(q13;q32). Two technologists independently scored 250 nuclei for each specimen in a blinded fashion; thus, 500 nuclei were scored for each specimen. If necessary, scoring criteria are modified on the basis of results obtained in Experiment 2. Because the analysis of the data is similar to Experiment 2, the Table 2 format can be used to document the data for each patient in Experiment 3. In this experiment, 89% of all cells met the scoring criteria. We encountered three new signal patterns that were not in the scoring criteria, namely, 2R1G2F, 1R2G2F, and 1R1G3F; each of these signal patterns was added to the final scoring criteria. The maximum number of false-positive nuclei was 8 with 2R3G in any single patient, 35 with 1R1G1F in any single patient, and 0 for all other signal patterns. On the basis of these figures and the beta inverse function, the normal cutoff was set at 2.83% for 2R3G, 8.58% for 1R1G1F, and 0.60% for all other patterns. The reference range for abnormal specimens was also determined for each signal pattern as shown in Table 3. For example, the abnormal reference range was 37% to 64% for 2R3G and 12% to 93% for 1R1G2F.

Experiment 4: Precision

For metaphase and qualitative interphase FISH assays, the analytic sensitivity and specificity are useful estimates of the precision of the FISH assay. For quantitative FISH assays, it is necessary to perform more elaborate experiments such as proposed here. It is useful to select a specimen for this experiment that can also serve as a standard control for clinical practice.²²

Precision

Precision is the ability to obtain the same result on each run and on each day. The standard deviation may be small or great depending on random variations, such as sampling error. Precision is also a useful tool for monitoring the procedure over time.

DISCUSSION

Validation is required to comply with regulatory standards. Moreover, validation is important to verify that testing will be consistent, safe, and accurate before use in clinical practice. The validation process can be confusing because of the lack of step-by-step procedures. This report should provide the required elements of the validation process.

For rare disorders, some laboratory directors are challenged with obtaining suitable normal and abnormal specimens to perform validation studies. These individuals may want to collaborate with other investigators or obtain abnormal cell lines from mutant cell repositories. An excellent method to accumulate samples for validation is to bank residual cell pellets from clinical practice.

Procedures to validate FISH assays vary depending on specific clinical applications and FISH strategies, whether the probes are approved by a regulatory agency, and the extent of published work regarding the test. Only a pilot study is needed if the FDA or similar regulatory authority has validated and approved a FISH assay and the procedure is followed as written by the manufacturer. The assay must be used only for the clinical application for which the regulatory agency has approved the product.¹⁶ If the FDA-approved method or clinical application is not followed exactly, it is necessary to validate the procedure and establish performance criteria before testing patients in clinical practice.

The ability of any technologist who performs the new assay in clinical practice must be evaluated before he or she does clinical testing. Technologists' competency at scoring FISH signals can be assessed on a routine basis by evaluating the interobserver variation for the assay, and the consistency of applying the scoring criteria as measured over time.²² If this difference is significant, the source of variation should be investigated and documented.

In clinical practice, unexpected signal patterns may be observed because of chromosome abnormalities that were not encountered in either the pilot study or clinical evaluation. If a novel abnormal signal pattern is present, it should be verified on metaphase cells or at least correlated with the karyotype result or clinical presentation.

Control specimens may or may not be used during the preclinical validation process. During validation, control specimens are not required because known normal and abnormal specimens serve as their own controls. However, controls are required in clinical practice to detect variability because of problems with reagents, equipment, technologist performance, and other factors. CLIA standards require the use of control specimens in laboratory testing, including FISH studies, to detect "immediate errors" in all steps of a single test as well as "long-term changes" in the testing system.³ FISH tests should include controls, internal or external, designed to detect errors, assess performance of the FISH test, and ensure the accuracy of scoring criteria. The reader may refer to Stupca et al.²² for more information on the use of controls in clinical practice.

This publication focuses on the preclinical validation process and is best suited for conventional FISH studies for metaphase and interphase cells. Nevertheless, the elements of the validation process would also be applicable to emerging molecular cytogenetic technologies such as array CGH analyses.

ACKNOWLEDGMENTS

We thank Heather Flynn, Ryan Knudson, Sarah Paternoster, and Stephanie Smoley for their critical review.

References

1. Food and Drug Administration. *Guidelines on General Principle of Process Validation*. www.fda.gov/cder/guidance/pv.htm. Accessed 29 December 2004.
2. College of American Pathologists Laboratory Accreditation Checklists. http://www.cap.org/apps/docs/laboratory_accreditation/checklists/checklistftp.html. Accessed 29 December 2004.
3. Clinical Laboratory Improvement Amendments. *CLIA:42 CFR Part 493-1253*. <http://www.phppo.cdc.gov/clia/regs/toc.aspx>. Accessed 29 December 2004.
4. New York State Health Department Clinical Laboratory Evaluation Program. <http://www.wadsworth.org/labcert/lep/Survey/standardsmenu.htm>. Accessed 4 January 2005.
5. American College of Medical Genetics. *Standard and Guidelines for Clinical Genetic Laboratories*. <http://www.acmg.net>. Accessed 29 December 2004.
6. American College of Medical Genetics. Technical and clinical assessment of fluorescence in situ hybridization: an ACMG/ASHG position statement. I. Technical considerations. *Genet Med* 2000;2(6):356–361.
7. Schad CR, Dewald GW. Building a new clinical test for fluorescence in situ hybridization. *Applied Cytogenet* 1995;21:1–4.
8. Cohen N, Novikov I, Hardan I et al. Standardization criteria for the detection of BCR/ABL fusion in interphase nuclei of chronic myelogenous leukemia patients by fluorescence in situ hybridization. *Cancer Genet Cytogenet* 2000;123(2):102–108.
9. Hausmann M, Cremer C. Standardization of FISH-procedures: summary of the first discussion workshop. *Anal Cell Pathol* 2003;25(4):201–205.
10. Bilous M, Dowsett M, Hanna W et al. Current perspectives on HER2 testing: a review of national testing guidelines. *Mod Pathol* 2003;16(2):173–182.
11. Zarbo RJ, Hammond ME. Conference summary, Strategic Science symposium. Her-2/neu testing of breast cancer patients in clinical practice. *Arch Pathol Lab Med* 2003;127(5):549–553.
12. Dewald GW, Stallard R, Bader PI et al. Toward quality assurance for metaphase FISH: a multicenter experience. *Am J Med Genet* 1996;65(3):190–196.
13. Dewald G, Stallard R, Al Saadi A et al. A multicenter investigation with interphase fluorescence in situ hybridization using X- and Y-chromosome probes. *Am J Med Genet* 1998;76(4):318–326.
14. Dewald G, Stallard R, Alsaadi A et al. A multicenter investigation with D-FISH BCR/ABL1 probes. *Cancer Genet Cytogenet* 2000;116(2):97–104.
15. Dewald G. Interphase FISH studies for chronic myeloid leukemia. In: *Methods in Molecular Biology: Molecular Cytogenetics: Protocols and Applications Vol 204*. Fan YS (Ed.) Totowa, NJ: Humana Press, 2002:311–342.
16. National Committee for Clinical Laboratory Standards. *Fluorescence In Situ Hybridization (FISH) Methods for Medical Genetics: Approved Guideline*. NCCLS Document MM7-A. [ISBN1-56238-524-0]. NCCLS, 940 West Valley Road, Suite 1400, Wayne, Pennsylvania 19087-1898 USA, 2004.
17. Montgomery K, Keitges ES, Meyne J. Molecular cytogenetics: definitions, clinical aspects, and protocols. In: *The AGT Cytogenetics Laboratory Manual 3rd ed*. Barch M, Knutsen T, Spurbeck JL (Eds.). Philadelphia: Lippincott-Raven, 1997:561–563.
18. Dewald G, Brockman SR, Paternoster SF. Molecular cytogenetic studies in hematological malignancies. In: *Hematopathology in Oncology*. Finn W, Peterson L (Eds.). Kluwer Academic Publications, 2004:69–112.
19. Ahmann GJ, Jalal SM, Juneau AL et al. A novel three-color, clone-specific fluorescence in situ hybridization procedure for monoclonal gammopathies. *Cancer Genet Cytogenet* 1998;101(1):7–11.
20. Halling KC. Vysis UroVysion for the detection of urothelial carcinoma. *Expert Rev Mol Diagn* 2003;3(4):507–519.
21. Thall P, Jacoby D, Zimmerman S. Estimating genomic category probabilities from fluorescent in situ hybridization counts with misclassification. *Appl Stat* 1996;45:431–446.
22. Stupca P, Meyer RG, Dewald GW. Using controls for molecular cytogenetic testing in clinical practice. *J Assoc Genet Tech* 2005;31:4–8.