

# **Guidance for Industry**

## **In the Manufacture and Clinical Evaluation of *In Vitro* Tests to Detect Nucleic Acid Sequences of Human Immunodeficiency Viruses Types 1 and 2**

Additional copies are available from:  
Office of Communication, Training and  
Manufacturers Assistance (HFM-40)  
1401 Rockville Pike, Rockville, MD 20852-1448  
(Tel) 1-800-835-4709 or 301-827-1800

(Internet) <http://www.fda.gov/cber/guidelines.htm>

**U. S. Department of Health and Human Services  
Food and Drug Administration  
Center for Biologics Evaluation and Research (CBER)  
December 1999**

# TABLE OF CONTENTS

Note: Page numbering may vary for documents distributed electronically.

<b>I. INTRODUCTION.....</b>	<b>1</b>
<b>II. INTENDED USE.....</b>	<b>3</b>
<b>III. MANUFACTURING.....</b>	<b>3</b>
A. Rationale and Design.....	3
B. Assay Optimization.....	4
C. Sample Preparation.....	5
D. Primers and Probes.....	5
E. Reaction Buffers.....	6
F. Enzymes.....	7
G. Controls and Calibrators.....	7
H. Other Test Kit Components.....	9
I. Detection and Quantitation of Amplicons.....	9
J. Instrumentation and Software.....	9
K. Sterility/Bioburden.....	10
L. Kit and Component Stability.....	10
<b>IV. CLINICAL VALIDATION OF ASSAY PERFORMANCE.....</b>	<b>10</b>
A. Preclinical Studies.....	10
B. Clinical Trials: General Issues.....	12
C. Specificity and Sensitivity Studies for Test Kits with a Proposed Labeling for Screening of Blood and Plasma Donors.....	14
D. Studies to Validate Intended Use as Additional, More Specific Tests.....	17
E. Clinical Prognosis and Management of Patients on Therapy.....	18
F. Perinatal Diagnosis.....	21
<b>V. CONCLUSIONS.....</b>	<b>21</b>
<b>VI. REFERENCES.....</b>	<b>22</b>

# GUIDANCE FOR INDUSTRY<sup>1</sup>

## In the Manufacture and Clinical Evaluation of *In Vitro* Tests to Detect Nucleic Acid Sequences of Human Immunodeficiency Viruses Types 1 and 2

### I. INTRODUCTION

In March 1985, the U.S. Food and Drug Administration (FDA) licensed the first screening test for the detection of antibodies to Human Immunodeficiency Virus (HIV) in serum and plasma from infected individuals. As of October 1999, there were 27 licensed kits and 5 premarket approvals (PMAs) for detection of antibodies to HIV-1 or HIV-2 in blood, saliva, or urine, which included: 19 Enzyme-Linked Immunosorbent Assays (ELISAs), 3 Western Blots, 1 Particle Agglutination Assay, and 1 Indirect Immunofluorescence Assay (IFA) for detection of antibodies to HIV-1, as well as 3 ELISAs for detection of HIV-1 p24 antigen, 2 of which are for use in donor screening. The ELISAs include 4 combination tests for detection of antibodies to HIV-1 and HIV-2 in blood specimens.

FDA recently licensed the first test for the detection of HIV-1 antibodies in urine specimens. HIV ELISAs have been approved to screen blood and plasma donors. The IFA, the Rapid Latex Agglutination Assay, and the colorimetric Single Use Diagnostic System (SUDS), may be used to screen blood donors in urgent situations. They are primarily used for urgent testing in hospitals, laboratories, medical clinics, physician's offices, emergency care situations, blood banks, or other health care settings when a routine ELISA is unavailable or impractical. Repeatedly reactive results from screening assays are further evaluated by additional, more specific tests which include Western Blot and IFA. In December 1994, FDA approved the first oral fluid collection device for professional use with a licensed HIV-1 antibody test kit and in June 1996, a supplemental Western Blot test to further evaluate the presence of antibodies in oral fluid was approved. Two PMAs for home blood sample collection kits, labeled as part of a HIV testing system, were also approved by FDA in 1996 (one has been voluntarily discontinued by the manufacturer). The first PMA for a HIV-1 nucleic acid test, a quantitative HIV-1 ribonucleic acid (RNA) test based on amplification of deoxyribonucleic acid (DNA)

---

<sup>1</sup> This guidance document represents the agency's current thinking on *in vitro* testing to detect specific nucleic acid sequences of HIV. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. An alternative approach may be used if such approach satisfies the requirements of the applicable statute, regulations, or both.

sequences using the polymerase chain reaction, to measure viral load in plasma as an aid in determining patient prognosis, was approved by FDA in June 1996.

In recent years, several technical advances have been made in methodologies for direct detection of viral nucleic acid. This document provides guidance on manufacturing and clinical trial design issues pertaining to the validation of tests based on nucleic acid detection either in the presence or absence of an amplification step. Concerns regarding the procedures used for detection of amplified products are also addressed. As may be the case with new technologies, issues may be identified during the review process, unique to the particular methodology under review or the specific configuration of the assay that will need to be addressed on a case-by-case basis. It is also recognized that this area of science is in a state of rapid technological development. As advances are made, this document will be reevaluated and revisions or modifications made as necessary. The criteria outlined below address both general and specific concerns for nucleic acid based detection techniques for HIV. This document is intended for products regulated by the Center for Biologics Evaluation and Research (CBER). The FDA uses mandatory language, such as shall, must, and require, when referring to statutory or regulatory requirements. The FDA uses non-mandatory language, such as should, may, can and recommend when referring to guidance.

The reader is referred to the Points to Consider in the Manufacture and Clinical Evaluation of *In Vitro* Tests to Detect Antibodies to the Human Immunodeficiency Virus Type 1 (1989) (Ref. 1) for general information on filing of the Investigational New Drug Applications (INDs), Product License Applications (PLAs), Establishment License Applications (ELAs), and content of applications for approval and licensure of retroviral kits. General regulations related to these products are located in 21 CFR parts 312, 600-680, and 800. Other documents that may be pertinent to this topic include the "Review Criteria" document (Ref. 2) issued by the Center for Devices and Radiological Health (CDRH), and the Guidance for Industry: Content and Format of Chemistry, Manufacturing and Controls Information and Establishment Description Information for a Biological *In Vitro* Diagnostic Product (Ref. 3) issued by CBER.

As set forth in the Intercenter Agreement of 1991 between CBER and CDRH, *in vitro* tests for HIV, that are recommended for blood donor screening and related blood bank practices, are licensed under the Public Health Service Act (PHS Act) through the IND/PLA/ELA/Biologics License Application (BLA) mechanism. *In vitro* tests for HIV, that are not performed in relation to blood bank practices (e.g., quantitative HIV assays and diagnostic tests that evaluate specimens other than blood), will be regulated by CBER under the Medical Device Authorities through the Investigational Device Exemption (IDE)/Premarket Approval Application (PMA) mechanism. *In vitro* diagnostics for pathogens other than HIV are regulated by CDRH under the IDE/PMA mechanism.

The scientific and regulatory concerns pertaining to validation of *in vitro* diagnostic/screening test kits can be broadly classified into three main categories: 1) intended use; 2) manufacturing; and 3) clinical validation of assay performance.

## **II. INTENDED USE**

The sponsor should state clearly in the application (IND, PLA/ELA, BLA, or IDE/PMA) the intended use, the labeling claims, and the clinical utility for the product. The proposed clinical trial design should be capable of demonstrating assay performance at a level that is sufficient to validate the intended use claim in the target patient population and the specific test setting. It is recommended that the sponsors/manufacturers meet with CBER to obtain guidance early in the development process in order to resolve any issues with regard to an approvable claim for the product or special concerns related to the product, and to address any questions the manufacturers may have. This should include a discussion of the proposed claim for clinical utility and the clinical studies that will be performed to validate the proposed claim, including equivalence or superiority to existing methods or licensed tests, if available, for detection and/or quantitation of the same agent.

## **III. MANUFACTURING**

The manufacturing issues that may have an impact on product design and performance are: 1) rationale of assay design; 2) assay optimization; 3) sample collection, extraction, storage and stability; 4) manufacture of primers, probes, reagent buffers, enzymes, calibrators, controls, and quantitation standards; 5) anchoring components, i.e., beads, plates, chips; 6) kit stability; and 7) instrumentation and software. This document addresses concerns pertaining to each of the points outlined above.

### **A. Rationale and Design**

The sponsor should provide in the application, the rationale for the specific indication and for use of the specific test methodology, and type of nucleic acid target (DNA or RNA) for detection of the infectious agent and for the specific indication. A detailed description of all aspects of the technique including sample preparation, assay optimization, amplification, and detection methods should be provided. Validated quality control procedures that are state-of-the-art should be used to assure manufacturing consistency. The sponsor should provide details of assay optimization and establish the format of the final product during the preclinical stage of development. Changes in assay format may lead to a recommendation for new studies. The rationale for assay design should address the following aspects:

1. Selection of target sequence(s) in the template including the degree of nucleic acid sequence conservation, guanine:cytosine (GC) ratio, and length. For products intended to detect more than one virus subtype or species (e.g., a “multiplex” design) define the number of target nucleic acid sequences and the rationale for their selection;
2. Assay format (e.g., sample type, conjugate, detector);
3. Selection of primer and probe sequences (e.g., degree of nucleic acid sequence conservation); and
4. Design and nature of the quantitation standards for a quantitative assay.

## **B. Assay Optimization**

This phase is critical to product development and can have a significant impact on product performance. The sponsor should address the various aspects of optimization for nucleic acid extraction, target sequence, amplification, detection, quantitation, and instrumentation for these processes, and set specifications for performance. The application should contain information on the details of:

1. The length, region, specificity, and efficiency of primer and/or capture sequence;
2. Methods of extraction, amplification, hybridization, detection, and quantitation;
3. Percent recovery of nucleic acid for the total assay and for each significant step in the process of sample preparation;
4. Optimization of reaction conditions and kinetics of amplification with multiple primers or hybridization with multiple probes or both (e.g., for a multiplex format);
5. Internal and external assay calibrators/controls; and
6. Procedures to prevent cross-contamination.

During this phase of assay optimization, the sponsor should determine and define the optimal assay conditions, reaction kinetics and the lower bounds of reliable assay performance. For qualitative assays the assay cutoff, or reporting threshold, is the lowest concentration of HIV RNA copies per ml that the assay can reliably distinguish from HIV negative samples ( $\geq 95\%$  detection rate). For quantitative assays, the lower limit of reliable assay performance may be defined by two potentially distinct values; the lower limit of detection (LOD) and the lower limit of quantitation (LOQ). For the purpose of this document the LOD is defined as the lowest concentration of analyte that can be distinguished from a negative specimen with a predefined level of assay sensitivity. The LOQ is the lowest concentration of analyte that is distinguishable from a negative specimen with the same degree of sensitivity as the LOD that is also quantifiable with an acceptable degree of precision and accuracy (e.g., CV of  $\leq 35\%$ ).

The assay cutoff/reporting threshold or LOD should be well defined in terms of copy numbers and the unit of sampling. This limit should be validated by an established form of independent characterization (e.g., an (approved) amplification technique or a combination of more direct measurements such as particle counts, EM scanning and quantitation of RNA by optical density). The assay cut-off or LOD can be established by in-house testing then further defined or verified based on the data from clinical trials. If the assay is quantitative, additional studies should be conducted to examine:

7. Linearity in the readable range in order to ensure accurate interpolation of unknown specimens. This range should also be clinically meaningful to demonstrate the clinical utility of quantitation; and
8. Accuracy and reproducibility based on quantitation of analytical specimens on a standard curve (analytical sensitivity). These studies should also provide a preliminary estimate of the LOQ.

### **C. Sample Preparation**

The sponsor should specify the type of specimen (e.g., cells, plasma, whole blood, dried blood spots) and the template for amplification (DNA and RNA) and hybridization, as appropriate. The composition of the buffers, reagents, and detergent or chaotropic agents used for nucleic acid extraction should be clearly specified. The effect of anticoagulants and any potential inhibitors present in the sample or extraction buffers on assay performance should be evaluated.

Controls that monitor the efficiency of the extraction and reverse transcription (when the template is RNA) procedures should be included and whenever possible, these controls should simulate the actual sample type. Spiked controls are also acceptable.

The reproducibility of the sample preparation method should be determined under the specimen processing conditions including sample handling, storage, and shipping conditions. The sponsor should also verify possible interference of specimen processing reagents with reverse transcription, amplification, hybridization, detection, and quantitation. For pool testing, sample stability during pooling and the subsequent processing steps should be determined.

### **D. Primers and Probes**

The primers and probes are the main components of a nucleic acid based detection system and the performance of the assay is highly dependent on the quality of these reagents. The sponsor should provide the:

1. Rationale for selection of primers and probes including specific sequences used;

2. Justifications for alignments made to generate consensus sequences or best-fit modifications made to existent sequences, e.g., to permit maximum homology to several strains; and
3. Information on size, GC content, melting temperatures, hairpin or other secondary structures if any, and the nucleotide position on the genome map of the primers and probes.

For assays designed to detect or quantitate multiple HIV subtypes or variants, data should be provided to demonstrate that the primers and/or probes chosen are effective for all of the subtypes or variants identified in the label.

If synthetic oligonucleotides are used as primers and probes, details of the manufacture and purification should be provided. In addition, the following information should also be included:

4. The yield and composition for the first 3 lots (at a minimum) by absorbance and DNA fingerprinting, restriction endonuclease mapping or nucleotide sequence analysis;
5. A description of the chemical nature of the modification, for modified oligonucleotides and procedure(s) to insure lot to lot consistency of ligand content;
6. Nucleotide sequence analysis to establish the fidelity of the procedure for oligonucleotide synthesis;
7. The purity of the final product should be analyzed by an appropriate state-of-the art analytical technique (e.g., reverse phase high performance liquid chromatography, electrophoresis or ion exchange HPLC), that has been validated according to ICH guidelines;
8. Potency of primers and probes. This may be addressed by dilutional analysis comparing lot-to-lot consistency in functional efficiency or other methodology appropriate to the technology under development.

The analyses listed in 6-8 should be conducted on each lot of oligonucleotide manufactured as a routine part of new product development and characterization. If a high degree of consistency is demonstrated over time the sponsor may request a reduction in the frequency of required monitoring.

#### **E. Reaction Buffers**

The sponsor should demonstrate the identity and purity of reagents used in the preparation of reaction buffers that are employed in amplification, hybridization, and detection reactions. The potency and stability of the reagents on storage and under cycling conditions should be verified.



If reagents are obtained from vendors, the quality system regulations (21 CFR 820.50 and 820.80) require the manufacturer to establish and maintain procedures to ensure that all received reagents conform to specified requirements. The extent of control necessary will be related to the nature of the reagent, taking into account the effect of the reagent on the finished product. A certificate of analysis should be provided for purposes of verification and the criteria used for acceptance/rejection of specific reagents defined. If deviations from component specifications could result in the product being unsuitable for use, a sponsor may be expected to sample and test components.

## **F. Enzymes**

The source and function of all enzymes used in the assay should be identified and clearly defined. The identity, purity, potency, and specific activity should be demonstrated and criteria for acceptance established.

For rDNA-derived enzymes manufactured by the sponsor, the master and working cell banks should be characterized for cell and genetic stability, and freedom from adventitious agents. Plasmid stability should be monitored by assays that include restriction mapping or DNA sequencing. If restriction mapping is used for plasmid monitoring confirmation of enzyme amino acid composition and sequence by peptide mapping and amino acid sequencing should also be considered.

Enzyme preparations should be tested for other enzymatic activities, e.g., exonucleases and DNA and RNA dependent polymerase activities and specifications should be established. For enzymes obtained from vendors, the certificate of analysis should be provided. In addition, functional testing designed to assure that the component is suitable for its intended use, should be performed as part of establishing the acceptance criteria.

## **G. Controls and Calibrators**

Controls are important tools that allow the operator to verify that the assay has performed within accepted specifications and are, therefore, a vital component of any test kit. Controls should be separate from, and in addition to, reagents used to estimate the concentration of an unknown sample (i.e., standards or calibration reagents).

In nucleic acid analysis, there are several steps in the testing process, as outlined above, that should be monitored and verified. It is therefore advisable to include multiple controls or controls that serve multiple purposes in the final kit. The controls should reflect the specific

technology under development but will typically allow for monitoring of ultracentrifugation, extraction, amplification, hybridization, quantitation, contamination, etc. These controls should be similar to the specimen type whenever feasible although spiked controls may be acceptable, particularly for labile analytes.

Sponsors are strongly encouraged to include a minimum of two positive controls to monitor assay performance. A control at or near the LOD or assay cutoff/reporting threshold should be incorporated into any assay that will be read in a qualitative fashion. For the validation of individual assay runs with a diagnostic assay, it is recommended that this control be within three standard deviations of the assay cut-off/reporting threshold.

For quantitative assays the low concentration positive HIV RNA control for validation of individual runs should be within 3 standard deviations of the LOQ. The second positive RNA control may fall anywhere within the linear range of the assay. In the event that the assay LOD is different from the lower limit of quantitation sponsors are strongly encouraged to include an additional control at the LOD to allow laboratories to monitor, on a routine basis, their ability to detect RNA at that level.

Assays that have or are seeking a label claim for quantitation of multiple viral subtypes should make a subtype specific positive RNA control available for each subtype.

Multiple negative controls should be included such as non-target sequences and nucleic acid free controls to monitor for false positives resulting from contamination. Due to the high sensitivity of amplification assays, it is highly recommended that sponsors include control measures for prevention of contamination events.

Specifications for both positive and negative controls should be provided, as well as validation data supporting the proposed assay cut-off/reporting threshold value or LOD of the assay. The sponsor should define the source of the controls and calibrators, and have a plan for their continued renewal. Controls should be non-infectious, and validation of viral inactivation should be provided.

For quantitative assays, validation data should be provided for all quantitation standards and calibrators. Specifications and acceptance criteria should be established for each control/calibrator and for the collective set of controls. Quantitation should be based on co-amplification of a heterologous internal control and/or a competitive RNA template or co-hybridization, as indicated by the technology under development. For RNA assays, the efficiency of reverse transcription should be determined for the specific assay format.

## **H. Other Test Kit Components**

The sponsor should provide a description of the anchoring solid phase component (e.g., plates, beads, filters), concentration of antigen or oligonucleotide on the component method of conjugation, or binding to the component, and a demonstration of lot-to-lot consistency of manufacture of bound component.

If more than one component is used for coating (e.g., two oligonucleotides) a description of the validation of coating methods including ratios used and acceptance criteria for the coating process should be provided. If the sponsor purchases a solid phase component (e.g., beads, plates, chips) a description of the source, quality assurance methods, and acceptance criteria should be included.

## **I. Detection and Quantitation of Amplicons**

A detailed description of the chemical/biochemical nature of capture probes, conjugates, detectors, quantitation standards, etc., which are part of the assay system should be provided. This should include:

1. The chemistry and limits of detection of system of choice (e.g., chemiluminescence, fluorescence);
2. Chemical and biochemical characterization of the ligand, chromophore, fluorochrome, including stability under reaction conditions;
3. Quality control and assurance of conjugation to detect or capture oligo- or polynucleotide sequence, including functional testing; and
4. Nature and copy numbers of the quantitation controls and standards.

Validation data should be provided for controls and quantitation standards. Specifications should be established for the individual and/or collective set of controls/quantitation standards used to detect/quantitate nucleic acid.

## **J. Instrumentation and Software**

Any dedicated equipment used in the amplification, detection, and quantitation of the amplified product should be validated for its use. These may include devices such as thermal cyclers, waterbaths, luminometers and cycling ovens.

Validation of thermal cyclers should include demonstration of the accuracy of temperatures of individual wells during the cycling process, specify limits for well-to-well variation, if any, as well

as any impact there may be on test results. If software is utilized for amplification, detection, and calculation of quantitative or qualitative results, validation of such software for the intended function should be provided.

For non-dedicated instruments, the premarket notification (510k) submission number should be cited for review. If previously approved under a PMA, a supplement for use with the product under review should be submitted.

If special specimen collection, storage and/or transport devices are used, specifications should be provided for conditions of collection, storage, and transport. Criteria should be established for suitability and adequacy of the specimen for the test.

#### **K. Sterility/Bioburden**

Refer to the "Points to Consider in the Manufacture and Clinical Evaluation of In Vitro Tests to Detect Antibodies to Human Immunodeficiency Virus Type 1" (1989) for guidance in this area.

#### **L. Kit and Component Stability**

The stability of the final kit and individual components should be tested using a panel of specimens, including weak reactives (e.g., near cut-off, middle, and upper end of the readable range). A reference panel consisting of plasma spiked with known amounts of virus is acceptable for this study. The real time stability at storage and shipping temperatures should be evaluated using specimens with varied reactivities in the readable range.

### **IV. CLINICAL VALIDATION OF ASSAY PERFORMANCE**

#### **A. Preclinical Studies**

Preclinical studies, performed either in-house or at field sites, provide preliminary information on assay performance. These studies should be designed to assess the sensitivity, specificity, and reproducibility of the test kit, as well as to identify the lower bounds of reliable assay performance (i.e., the assay cut-off/reporting threshold, LOD and LOQ, as appropriate for the assay under development). In general, preclinical testing should be performed prior to initiation of clinical validation studies, particularly if prospective, large scale clinical trials are planned. Clinical validation of these preliminary assessments is then accomplished through field testing of

appropriate clinical specimens to provide final specifications for test kit performance characteristics.

1. Specificity and Sensitivity Studies for Preclinical Testing (statistical determination of false positive and false negative rates)
  - a. For donor screening assays or diagnostics, specificity should be established by testing samples from a minimum of 500 random blood or plasma donors. For quantitative assays, a minimum of 100 samples should be tested in the preclinical development phase; and
  - b. For both assay types, sensitivity should be established by testing at least 300 seropositive repository specimens.

Testing should be performed in parallel with appropriate licensed comparator assays (i.e., antigen or antibody assays for blood screening or diagnosis and a nucleic acid based test for quantitation).

## 2. Analytical Sensitivity

Analytical sensitivity should be evaluated by testing a dilution series of at least 10 distinct HIV positive samples obtained from different individuals (clinical specimens). Prior to initiation of these studies, the starting concentration of viral RNA in each sample should be determined by an appropriate independent technology (see section III.B, above). The evaluation of patient isolates should be run in parallel with samples from a recognized reference panel (e.g., reference HIV panels produced by the WHO, CBER or the ACTG VQA).

For quantitative assays the dilution series should cover the full range of accurate quantitation for the assay under development. For pool testing, the lowest concentration in the dilution series must be at or below the minimum sensitivity level required for these assays (i.e., 100 copies/ml in a plasma pool or 5,000 copies/ml in an individual sample, based on a 95% detection rate). The highest dilution reproducibly and consistently detected and/or quantitated by the investigational assay should be defined in copies per unit sampling until recognized international standards for viral quantitation are available for the target virus, at which time, analytical sensitivity can be expressed in international units/ml. Comparator assays should include an antibody, antigen, or other state-of-the-art amplification/probe

technology. The selection of an appropriate comparator assay will be based on the technology under development and the proposed indication.

For products seeking a labeling claim for quantitation of HIV that includes non-clade B viruses, analytical sensitivity should be demonstrated for each subtype or variant proposed for inclusion in the label. A minimum of 10 distinct HIV positive specimens obtained from different individuals would be advisable for each subtype/variant. In the event that 10 distinct clinical specimens of a specific viral subtype are not available, culture derived specimens may be used to supplement clinical samples.

## **B. Clinical Trials: General Issues**

Clinical trials designed to assess clinical sensitivity, specificity, and reproducibility should be performed at clinical trial sites by qualified independent investigators. Refer to section VI. in the "Points to Consider in the Manufacture and Clinical Evaluation of In Vitro Tests to Detect Antibodies to the Human Immunodeficiency Virus Type 1" (1989) for additional general guidance on clinical trial design issues.

Common components of a clinical development program for any specific intended use include studies designed to assess the precision, reproducibility, and non-specificity of the investigational assay.

### **1. Precision studies**

These studies are designed to assess the coefficient of variation for the test results for each sample and for the various lots tested.

#### **a. Proficiency Testing**

A panel of samples similar to that described in section IV.B.1.b (below) should be tested by a number of operators on multiple days, at all clinical testing sites. This study is designed to assess operator proficiency.

#### **b. Reproducibility and Precision**

A panel of 10 or more samples including low reactives should be tested at all clinical trial sites. This could be a series of samples spiked with the analyte or human specimens with known reactivity. Samples should be tested in duplicate or triplicate. Testing should be performed on a minimum of three different lots, on multiple days and by at least two operators. The

operators chosen to conduct these studies should have demonstrated a high degree of proficiency with the assay. Reproducibility studies should be designed to assess variability intra- and inter-site, intra- and inter-assay and intra- and inter-lot, as well as total variability for both qualitative and quantitative assays. Assay precision may be established by performing multiple tests using multiple operators and multiple kit lots on a panel of specimens. Testing may be performed in-house and at least one clinical site.

c. Instrumentation

Instrumentation effects on product performance should be evaluated using the sample panel employed for reproducibility testing and a minimum of three different machines.

2. Non-specificity studies

Most assays are subject to some inherent non-specific reactivity resulting in false positive reactions or interference resulting in false negative reactions or reduced accuracy of quantitation. This effect may be due to specific assay components or the nature of the sample being tested. The presence of non-specific reactivity and the impact of potential interfering factors on assay performance should be assessed. Appropriate samples for these studies can be obtained by spiking the agent or factor into known HIV positive and/or negative samples, in addition to testing original specimens or reference panels. Examples of conditions, factors or sample characteristics that should be considered in the evaluation of cross-reactivity or interference include the following:

- a. Other infections including Human T-cell Lymphotropic Virus Type I/II (HTLV-I/II), Cytomegalovirus (CMV), Epstein-Barr Virus (EBV), Hepatitis B Virus (HBV), Hepatitis C Virus (HCV), yeast infections, and pneumocystis;
- b. The impact of different anticoagulants (i.e., heparin, ACD, or EDTA) or other suitable collection tube/media (e.g., plasma preparation tube);
- c. Hemolyzed, icteric, lipemic, and bacterially contaminated samples;
- d. Samples treated with chemicals, drugs, heat or detergents;
- e. Samples subjected to multiple freeze thaw cycles;
- f. Fresh vs. frozen samples, serum vs. plasma, and single specimen vs. plasma pool;

- g. Samples from patients with autoimmune diseases including Systemic Lupus Erythematosus (SLE), Anti-Nuclear Antibodies (ANA), and Rheumatoid Arthritis;
- h. The impact of subject age, gender, race, or ethnic group;
- i. The impact of drug resistance mutations. This evaluation should include a cross section of viruses that demonstrate phenotypic and genotypic resistance to each of the currently approved drug classes. Multiple drug resistant viruses should also be evaluated;
- j. The presence of nucleic acid based drugs and metabolites and binding substances; and
- k. The presence of drugs or biologicals that increase circulating nucleic acid.

### **C. Specificity and Sensitivity Studies for Test Kits with a Proposed Labeling for Screening of Blood and Plasma Donors**

FDA does not currently envision discontinuation of antibody testing for donor screening. However, nucleic acid testing (NAT) may be more sensitive than other methods currently available for early detection of a virus during the pre-seroconversion phase of infection and may, therefore, have added value in blood safety. For NAT, clinical specificity should be established by testing a large number of specimens from random U.S. blood and plasma donors. The evaluation of clinical specificity should include samples from at least 10,000 individuals or plasma pools, based on the intended use of the product. Testing should be performed at a minimum of 3 distinct clinical sites, including areas of both high and low prevalence. Studies should be performed in a manner that allows for donor identification (i.e., “linked”) to permit clinical follow-up. If positive results are encountered, the donor should be deferred temporarily and units held in quarantine until test results are confirmed by repeat NAT testing and/or by clinical follow-up and antigen or antibody-based testing over the subsequent 3-6 month period. The basis for reinstatement of donors with false-positive investigational test results should be defined in the study protocol.

#### **1. Additional Studies to Establish the Clinical Sensitivity of Tests Intended for Screening of Individual Blood Donations**

A minimum of 1,000 specimens from seropositive individuals including samples from various risk groups and different stages of HIV-1 disease should be tested at 3 distinct sites. At least 200 of these seropositive specimens should be derived from persons with a clinical diagnosis of AIDS. It is recommended that the study also include a gender-



based analysis with at least 20-30% of the samples derived from females.

Geographically diverse specimens representing all known viral subtypes (a minimum of 20 of each) should be included in this data set to evaluate the performance and establish the sensitivity of the kit for detection of variant HIV-1 strains. All of these samples should be actual clinical specimens as opposed to culture derived virus stocks or cloned species. However, if 20 distinct clinical specimens of a specific viral subtype are not available, culture derived specimens may be used to supplement clinical samples.

A minimum of 200 samples known to be positive for HIV-2 should also be tested in the evaluation of clinical sensitivity for single donation screening tests, including a subset that contain both HIV-1 and -2. These samples may be obtained from a repository.

Prospectively collected, freshly drawn specimens from individuals at high risk for HIV-1 should be tested in a linked study so that a minimum of 50 positive cases are identified (by serology, p24 antigen or NAT) to estimate sensitivity. A similar prospective trial in a population at high risk for HIV-2 should also be conducted, with sufficient enrollment to identify a minimum of 30 seropositive cases. These studies should include methods to objectively resolve discordances between investigational and comparator assays that include both antibody and antigen-based tests. This may require follow-up testing of the study subjects. A subset of these samples (100) should be tested by another state-of-the-art test (amplification/probe test).

A comparison should be made between freshly drawn and frozen specimens, as well as paired serum and plasma or other specimens (e.g., Whole Blood). The purpose of this study is to establish the comparability of the two storage conditions and the two specimen types, respectively.

## 2. Additional Studies to Establish the Clinical Sensitivity of Assays Intended for Pool Testing

Clinical sensitivity testing for pool tests should include the evaluation of 1,000 known seropositive specimens in a minimum of 100 separate pools. Of these, at least 25 pools should contain weakly reactive seropositive specimens introduced singly into the pool. All pools should contain at least several known negative specimens. The positions of the seropositive specimens should vary from one pool to another, in a random fashion. These studies should be performed at 3 sites, one of which could be in-house. In addition, the ability to detect known HIV subtypes/variants and HIV-2 should be determined using a minimum of 20 samples for each variant. These samples may be

plasma pools spiked with the appropriate subtype/variant or a virus preparation diluted in seronegative plasma. Additional HIV-2 samples may be required to support a label indication for detection of HIV-2. These studies should demonstrate that the NAT is capable of detecting a minimum of 100 RNA copies/ml in the pooled sample or 5,000 copies/ml of virus in the original donation, with a  $\geq 95\%$  detection rate.

### 3. Additional Issues for NAT Intended for Screening of Plasma Pools

If testing will be done on pooled specimens, additional issues to be addressed include:

- a. Demonstration of equivalent or superior sensitivity of the assay for testing donor pools compared to currently licensed methods for donor screening by laboratory and field testing;
- b. Rationale for the proposed pool size;
- c. The impact, if any, of possible interference or matrix effects generated during pooling, on test performance;
- d. Sample stability during collection, storage, pooling and transport, including a comparison of assay performance and sample stability with freshly drawn versus frozen specimens;
- e. Procedures for logging and tracking of specimens in a given pool, including validation of the pooling process (i.e., procedures included to verify that appropriate test samples are obtained for all donations included in a pool and that all donations intended for a specific pool are actually included);
- f. Specimen retrieval procedures to identify a positive specimen in a positive pool;
- g. Quality assurance in computing and reporting test results;
- h. Validation of instrumentation and automation; and
- i. Validation of software that may be used in conjunction with any of the procedures listed above.

Testing of specimens from appropriate primate models may also be useful towards establishing utility for detection of infection in the pre-seroconversion phase.

### 4. Clinical Validation of Assay Performance for Blood Screening Tests

Because of the limited availability of specimens in the antibody negative, pre-seroconversion phase, the agency urges the use of specimens from the following categories for clinical validation of assay performance for all blood screening tests: ongoing cohort studies and retrospective investigations; as well as a large number of seroconversion panels and specimens from high risk individuals enrolled in

prospective studies being conducted in areas of high prevalence. These data should be analyzed on the basis of mathematical models that estimate the timing and duration of the window period. Testing should include a head-to-head comparison with a licensed assay for HIV p24 antigen. These studies are particularly important for NAT, where the value added may be in the ability to reliably detect early infections.

#### **D. Studies to Validate Intended Use as Additional, More Specific Tests**

These tests are used to further evaluate the accuracy of the positive test results of a screening assay. Gene based tests may be developed as an alternative to Western Blot (WB), Strip Immunoblot Assay (SIA), or Immunofluorescence Assays (IFA) currently in use for this purpose. In some instances, the test may be used to resolve the indeterminate patterns seen on additional, more specific tests including those referred to above.

In specificity and sensitivity testing for confirmatory tests, random donors should be tested at two or more sites. At least 500 samples should be tested to assess specificity. Approximately 300 specimens from random donors that are repeatedly reactive (RR) by licensed screening assays should be evaluated along with other additional, more specific tests to establish clinical sensitivity. In addition, a minimum of 300 known positive specimens should be tested to determine clinical sensitivity.

For tests that are used to resolve indeterminate results of other confirmatory assays, a minimum of 300 samples from persons with such indeterminate test results using licensed tests should be tested. Testing should be performed on a combination of WB, SIA, and IFA indeterminate specimens from random donors. Sponsors should demonstrate that the sensitivity of the second, more specific test is equivalent to or better than the screening assay.

In all cases, a plan for resolving discordant/discrepant results should be included. In addition, a minimum of 500 random donor specimens and 300 known positive specimens should be tested to establish clinical specificity and sensitivity.

## E. Clinical Prognosis and Management of Patients on Therapy

### 1. Clinical Specificity

Clinical specificity should be established by testing at least 500 specimens from healthy, random donors. This test series is distinct from the preclinical evaluation of specificity and should be conducted at an appropriate clinical trial site.

### 2. Clinical Sensitivity

Clinical sensitivity should be established by testing samples from seropositive individuals and from high risk groups in a head-to-head comparison with a licensed or approved nucleic acid detection test. Performance should be evaluated in cross-sectional studies involving HIV positive individuals stratified by CD4 counts and clinical history.

Category	No. of subjects
CD4 < 200	200
CD4 200-500	300
CD4 > 500	300

Studies designed to assess variability in consecutive viral nucleic acid measurements in an individual over time should also be performed. A minimum of 40 adult and 20 pediatric (age  $\leq 12$  years) patients with stable viral loads in the ranges listed above should be followed on a bi-weekly basis for a period of 16 weeks or weekly for a minimum of 8 weeks. The data should be evaluated with respect to short-term variability (1-2 weeks) and at 8 week intervals (typical monitoring frequency for routine patient care). The distribution of male and female participants in each adult study cohort should be sufficient to allow for the evaluation of gender based differences in assay variability.

The information obtained from the studies listed above will be important in establishing the significance and validity of changes observed after initiation of a specific therapy and the utility of these tests in patient management.

### 3. Performance in Patients Undergoing Therapy

The clinical utility of an assay in treated individuals should be demonstrated in prospectively conducted drug efficacy trials by direct comparison to an approved nucleic acid quantitation test. Randomized, concurrently controlled clinical trials that compare different drug combinations or investigate the utility of new drugs as adjuncts to existing therapies based on their effects on plasma virus RNA are acceptable for assay comparator studies.

A possible alternative approach to testing in the context of a prospectively conducted clinical study is the use of well-characterized, repository specimens from cohorts for which clinical outcome is known. However, if a retrospective analysis is planned, the study employed should be consistent with current treatment regimens and practices. In addition, it is essential that specimen adequacy requirements be confirmed prior to initiation of the study so that meaningful results may be evaluated and discordant results resolved. Prior to the conduct of a retrospective analysis of clinical utility, sponsors are strongly encouraged to consult with the division about the suitability of the study(s) selected.

Regardless of the approach taken to demonstrate assay utility in patient monitoring a minimum of 300 patients who underwent therapy with regular monitoring of virus load should be evaluated. This study population should include a subset of patients whose therapy was modified as a result of clinically defined disease progression, changes in virus load and/or the development of drug resistance. In the context of these studies, an effort should be made to address gender and age-related (i.e., pediatric versus adult) differences in kit performance.

Sponsors interested in pursuing a claim for patient monitoring are encouraged to consult the FDA Draft Guidance for Industry: Clinical Considerations for Accelerated and Traditional Approval of Antiretroviral Drugs Using Plasma HIV RNA Measurements (Ref. 4) for additional information on clinical testing strategies and accepted use of plasma RNA measurements in clinical therapeutic trials.

### 4. Clinical Prognosis

A number of different approaches may be taken to establish the prognostic value of a nucleic acid quantitation assay. A prospective study may be conducted involving

patients at different stages of disease with monitoring of their disease free time interval as the primary endpoint. Data may also be derived from well-characterized, retrospectively collected samples obtained from appropriate clinical endpoint trials. In some instances the prognostic utility of an assay may be demonstrated by direct comparison to another nucleic acid technology that is currently approved for prognosis.

When prognostic utility is defined by evaluation of clinical trial samples with clinical outcome data, whether prospectively or retrospectively identified, a statistical analysis should be performed to determine the relative predictive value of nucleic acid levels as it relates to disease progression. This analysis should be based on a combination of longitudinal and cross-sectional studies that includes information on clinical outcome. The cross-sectional study should employ a design similar to that described above for clinical sensitivity. For the longitudinal study, a total of approximately 500 patients at different disease stages based on the Centers for Disease Control and Prevention classification scheme should be followed. Clinical outcome may be defined as the time to first AIDS defining event or duration of disease free survival. The studies outlined above should include a sufficient number of female participants to allow for the evaluation of gender based differences in prognostic utility. An independent evaluation of pediatric patients is also strongly encouraged.

A comparative approach to demonstrate equivalent or superior performance may be acceptable in combination with supporting preclinical and clinical data for modified versions of currently approved assays, or for new tests that approximate, in a clearly definable manner, the RNA values generated with a nucleic acid technology approved for prognosis. A minimum of 150 clinical samples spanning the full linear range of both the experimental and reference assays should be evaluated. In addition, a set of sequential clinical specimens from 30 individuals with demonstrable changes in viral load (a minimum 1  $\log_{10}$  change), over time should be evaluated to permit a comparison of the assays based on individual readings and on sequential changes in those readings. A comparative approach is not sufficient for assays that do not demonstrate an acceptable degree of correlation with an approved assay based on both the individual readings and on sequential changes in those readings.

## **F. Perinatal Diagnosis**

Nucleic acid tests may be useful for early diagnosis of infection in infants born to seropositive mothers. The specificity of tests intended for this purpose should be established by testing at least 300 healthy (adult) donors and approximately 100-200 samples from infants born to healthy, seronegative mothers. Clinical sensitivity for early diagnosis should be evaluated by testing at least 200 infants born to seropositive mothers. Testing should include the first 0-6 weeks after birth and the results should be compared with licensed antibody and antigen tests. Long term follow-up may be needed to resolve test results in some cases.

## **V. CONCLUSIONS**

Gene based tests for viral agents are regulated by FDA either as biologics or devices. Although this document outlines some of the major regulatory and scientific issues concerning gene based tests for HIV-1 and HIV-2, these considerations may also be applicable to tests for other transfusion transmitted viruses including HCV, HBV, and HTLV-I and II. Sponsors are advised to consult with the agency (CBER or CDRH, based on the assay and indication for use) during the development phase so that product specific issues in manufacturing and clinical trial design may be addressed early in the validation phase.

## VI. REFERENCES

1. “Points to Consider in the Manufacture and Clinical Evaluation of In Vitro Tests to Detect Antibodies to the Human Immunodeficiency Virus Type 1” (1989).
2. “Review Criteria for Nucleic Acid Amplification-Based In Vitro Diagnostic Devices for Direct Detection of Infectious Microorganisms” (1993).
3. “Guidance for Industry: Content and Format of Chemistry, Manufacturing and Controls Information and Establishment Description Information for a Biological *In Vitro* Diagnostic Product” (1999).
4. Draft “Guidance for Industry: Clinical Considerations for Accelerated and Traditional Approval of Antiretroviral Drugs Using Plasma HIV RNA Measurements” (1999).