

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
Contract No. HHSN266200700023C  
Standard Operating Procedures

**SUBJECT:** **Reproducibility (Replication) Studies for Antigen Assays**

**PURPOSE:** Testing requirements for each assay submitted by a third party will vary based on the stage of development of the assay and the intended use of the data. This SOP is a broad and generic template for evaluating reproducibility and is to serve as a foundation for developing a more specific and individualized evaluation plan for each assay received from a third party. This SOP provides a template for:

Estimating the precision of an experimental antigen test and comparing the results with the third party's claims or literature reports.

Establishing / confirming the limit of detection for an experimental antigen test

Establishing / confirming the linear range of an experimental quantitative antigen test

Estimating the accuracy of an experimental test using an independent standard, animal model sample, or existing inventories of banked human samples and comparing the results with the third party's claims or literature reports.

**LEVEL:** Principal Investigator/designee  
Laboratory staff

**SUPPLIES/  
EQUIPMENT:** Policy and Procedure Manual  
Third Party's Assay / Instrumentation  
Third Party's Assay Software  
Third Party's Test Procedure  
AsTeC Consortium Laboratory Equipment  
Test samples  
Pipettes and pipette tips  
Computer  
Printer

**REQUIREMENTS:** A specific GLP compliant SOP outlining test procedures according to third party's recommendations will be written for each antigen detection test method accepted for evaluation. The analyses described below will be

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
Contract No. HHSN266200700023C  
Standard Operating Procedures

conducted for each test method under evaluation following the individual test's specific SOP testing procedures.

If multiple matrix specimen types (urine, serum, spinal fluid, whole blood, etc.) are to be assayed, separate analyses as described below for each specimen type will be performed.

A. An initial 5 day (minimum) familiarization period will be employed to reduce bias due to operator inexperience. During this time, the operator will become familiar with the test device and technology. Quality control material will be run during this "break-in" period to establish stable results. At the end of the familiarization period, an initial evaluation of repeatability will be conducted as outlined below. (EP5-A2)<sup>1</sup>

B. Three samples containing the test analyte will be supplied by the third party containing low, high, and "near medical decision point" levels of analyte. Twenty aliquots (or a complete "batch" if less than 20) of each of the three samples supplied by the third party will be assayed. Quality control samples, as recommended by the third party, will be included.

C. The mean, standard deviation, coefficient of variation, and the standard error of the mean for the test results will be calculated.

D. If a considerable discrepancy from the expected CV at each analyte level (as determined by the AsTeC Review Committee, for example, >15% of the CV submitted by the third party) is found, the testing may be performed a second time using three new samples. If considerable discrepancy is still encountered, the third party will be contacted and no further testing will be conducted until the problem is resolved.

E. This single run test is used to identify problems that should be resolved before continuing the evaluation of the test method. Results from this repeatability study will not be used as part of the reproducibility (replication) study.

F. If no considerable discrepancy from expected results is found, reproducibility (replication) studies will be conducted.

G. Precision

a. *For quantitative tests (EP5-A2)*<sup>1</sup>

To determine precision, three different studies (within-run, between-run, and between-day) will be conducted over 20 days employing a single reagent lot, calibration cycle, device, and operator. Three different samples with known amount of the test analyte will be used. The samples will be provided by the test third party, the Invasive

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
Contract No. HHSN266200700023C  
Standard Operating Procedures

Aspergillosis Animal Models (IAAM) contractor, or a sample will be prepared by the consortium laboratory by spiking an appropriate matrix with known amount of target or by serially diluting a sample with known high concentrations of the analyte. The 3 samples will have low, high, and “near medical decision point” (NMDP) levels of analyte.

Each sample will be analyzed a minimum of 2 times within a single analytic run (within-run) and 2 analytic runs will be performed each day (between-run) for a total of 4 replicates per specimen each day. The testing will be repeated on 20 separate testing days (between-day). To ensure acceptability of the run, quality control samples will be included in each run. If at any point an out-of-control condition is detected, the cause will be determined and the run repeated.

1. Recording Results

Results for each concentration of analyte will be recorded using the format of Data Sheet 1 in Appendix A. Quality control charts for the device will be set up at the end of the protocol familiarization period and all quality control data on the charts will be plotted. QC values  $>\pm 3$  SDs will be investigated and if an out-of-control condition is detected and resolved the run may be repeated. If the QC values of the repeated run fall within  $\pm 2$ SDs, the run can be accepted. If a QC value is  $\pm 4$  SDs the run will be rejected. A record of the number of rejected runs will be maintained.<sup>1</sup>

2. Statistical Analyses

*Separate calculations will be performed for each concentration of analyte.*

An initial test for within-run outliers will be performed. If the absolute value of the difference between a pair of replicates exceeds 5.5 times the SD determined in the preliminary precision test set (initial evaluation of repeatability), the pair of results within a run will be rejected. If the pair is rejected, the run will be repeated for that pair. If more than 5% of runs need to be rejected and no assignable cause can be found, the device may not be sufficiently stable to allow further assessment. The AsTec Review Committee will be consulted before proceeding with further testing.

For each run, day, and test cycle (days 1-20), the mean of the results will be calculated and the estimate of repeatability will be derived based on statistical

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
Contract No. HHSN266200700023C  
Standard Operating Procedures

calculations outlined in CLSI Document EP5-A2, Section 10.8 *Statistical Calculations for Precision*<sup>1</sup>. The standard deviation of the daily means as well as between-day and between-run standard deviations will be used to calculate the overall (total) precision of the device based on formulas presented in EP5-A2.<sup>1,6</sup> The coefficient of variation corresponding to the estimate of precision will also be calculated. Precision estimates will be compared to performance claims for the precision of the device. A chi-square test will be calculated as the ratio of the user's estimated within-run variance to the square of the performance claim standard deviation, multiplied by the degrees of freedom of the within-run variance estimate. If the test statistic exceeds the upper 95th percentile of the chi-square distribution with appropriate degrees of freedom, the null hypothesis that the estimate is not significantly different from the claimed value will be rejected.

b. *For qualitative tests with only two possible outcomes (EP12-A)*<sup>4</sup>  
(See Appendix B for definitions and further detailed explanations of cutoffs and intervals used in this section.) In order to estimate precision of a qualitative method at analyte concentrations near the cutoff, the third party will provide the 50% cutoff analyte concentration and the estimated lower limit of detection for the test. If the third party does not supply the 50% cutoff estimate, a dilution series should be made from a positive sample and dilutions tested in replicate to determine the dilution that yields 50% positive and 50% negative results. This dilution then contains the analyte concentration at the 50% cutoff point. Three samples will be prepared in sufficient volume to allow up to 20 replicate tests on each sample. Sample 1 will be prepared at the stated or determined 50% cutoff concentration, sample 2 will be prepared at a concentration 20% above the 50% cutoff concentration, and sample 3 will be prepared at a concentration 20% below the 50% cutoff concentration. Twenty replicates will be run on each of these 3 samples for a total of 60 results.<sup>4</sup> Results will be recorded using the format of Data Sheet 2 in Appendix A.

1. Statistical Analyses:  
The percent of positive and negative test results for each of the 3 samples will be calculated.

The sample at the cutoff should reveal a positive result ~50% of the time. If the 50% cutoff sample does not yield a

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
Contract No. HHSN266200700023C  
Standard Operating Procedures

positive result for ~50% of samples, the 50% cutoff concentration was inaccurate or the dose-response curve for the method is not linear near the cutoff point. If the 50% cutoff concentration was supplied by the third party, follow the procedure for calculating the 50% cutoff as outlined above, and repeat the experiment. If this experiment also fails to produce a positive result for ~50% of the samples, consult with the AsTeC Review Committee.

The sample prepared at 20% above the 50% cutoff concentration should be positive ~95% of the time and the sample prepared at 20% below the cutoff concentration should be negative ~95% of the time. If this is the case, then the 95% interval for the test is within  $\pm 20\%$  of the 50% cutoff and samples with analyte concentrations more than  $\pm 20\%$  of the 50% cutoff can be expected to yield consistent results with this test method and precision is acceptable.

If results for samples 20% above or 20% below the 50% cutoff fail to yield positive or negative results ~95% of the time, respectively, then samples more than  $\pm 20\%$  from the 50% cutoff cannot be expected to yield consistent results. In this case, precision is not established. Consult with the AsTeC Review Committee to determine if another series of experiments should be performed to determine the actual 95% interval for the assay.

#### H. Limit of Detection (EP17-A)<sup>9</sup>

The third party will provide the estimated lower limit of detection (LoD) and limit of the blank (LoB) for the test. (Refer to Appendix B for definitions of LoD and LoB.) Samples containing a low level of analyte and blank samples (identical to the test sample, but containing no analyte) will be used. Samples will be provided by the test third party, the Invasive Aspergillosis Animal Models (IAAM) contractor, or a sample will be prepared by the consortium laboratory by spiking an appropriate matrix with known amount of target or by serially diluting a sample with known high concentrations of the analyte.

Twenty to sixty replicates of a blank material will be performed (five measurements over at least four days). The number of replicates will depend on whether we are determining (60) or verifying (20) the claimed LoB. (If the LoB estimate is not available from the third party, the LoB will be determined

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
Contract No. HHSN266200700023C  
Standard Operating Procedures

following procedures described in CLSI EP17-A, page 12, Section 4.3.1.)<sup>9</sup>

Similarly, a sample will be prepared with concentration equal to the claimed LoD. Twenty to sixty replicates (five measurements over at least four days) will be run on the sample. (If the LoD estimate is not available from the third party, the LoD will be determined following procedures described in CLSI EP17-A, pages 12-13, Section 4.3.2.)<sup>9</sup>

1. Recording Results  
Results for each LoB and LoD replicate will be ranked according to size and will be recorded using the format of Data Sheet 3 in Appendix A.

2. Statistical Analyses:  
*To verify a Third Party Claimed LoD based on 20 replicates*

Percent of positive test results for the LoB sample will be calculated. If no more than 3 (15%) replicates on the blank exceed the claimed LoB, then the LoB is considered validated and can be used. If more than 3 replicates on the blank exceed the LoB, estimate the LoB as described in CLSI EP17-A, page 12, Section 4.3.1 and use this estimate in the following calculations.

The proportion of LoD sample results that exceed the LoB value will be determined. If the recorded proportion is in agreement with the expected value (95%) (i.e., if “95%” is contained within the 95% confidence limits for the recorded proportion), then the data support the claim of the LoD. Refer to CLSI EP17-A, Section 4.3.4, Table 2. *Lower 95% Confidence Bounds for Observed Proportions of Results Exceeding the LoB* to determine the lower bound of the 95% confidence interval based on the number of replicates tested.<sup>8</sup> In this experiment, where 20 replicates are performed, if the percent of LoD sample results that exceed the LoB is > 85%, the observed proportion is considered in accordance with the claimed LoD.

I. Linearity (for quantitative tests only) (EP6-A)<sup>7</sup>  
The third party will provide the estimated linear range for a quantitative test. The linearity of the quantitative assay will be evaluated to establish whether the test meets the third party’s

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
Contract No. HHSN266200700023C  
Standard Operating Procedures

linearity specifications using samples whose concentration levels are known relative to each other. These samples may be provided by the third party, the IAAM contractor, or a sample will be prepared by the consortium laboratory by spiking an appropriate matrix with known amount of target or by serially diluting a sample with known high concentrations of the analyte. The “seven sample” dilution scheme described in CLSI Document EP6-A, Appendix A, page 20 will be employed for this purpose.<sup>7</sup>

Seven to 11 levels of analyte will be tested including one sample with an analyte level 20% to 30% higher than the stated upper linearity limit and one sample with an analyte level at or below the stated lower linear limit. Each concentration level is run 4 times. All testing will be performed in a single day in closely grouped runs (or a single run when possible) with replicate specimens from each level tested randomly. Results will be recorded using the format of Data Sheet 4 in Appendix A.

1. Statistical Analysis:

Data will be plotted with analytical results on the Y axis, versus analyte concentrations on the X axis to visualize linearity and identify outliers. Single results that are visually different than other replicate values for that concentration are considered outliers. A single outlier in a dataset can be removed and does not need to be replaced. Two or more outliers casts doubt on the testing system’s performance and requires trouble shooting with the third party. Polynomial regression analyses will then be performed according to CLSI document EP6-A Section 5.3.2, page 11. A t-test will be performed to test whether the nonlinear coefficients are statistically significant ( $p < 0.05$ ). If none of the nonlinear coefficients are significant, the dataset is considered linear and the analysis is complete. If nonlinearity is detected, the AsTec Review Committee in consultation with the third party, will determine whether to proceed with additional testing to establish the degree of nonlinearity as described in CLSI Document EP6-A, section 5.3.3, page 12.<sup>7</sup>

- J. Accuracy (EP12-A, EP15-P, and FDA guidance documents)<sup>2,3,4,5,6</sup>  
Reference materials with analyte target values will be provided by the test third party, IAAM contractor, existing specimen banks containing samples from human subjects with proven IA (based on EORTC/MSG definitions), or a sample will be prepared by a consortium laboratory by spiking an appropriate matrix with

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
Contract No. HHSN266200700023C  
Standard Operating Procedures

known amount of target analyte or by serially diluting a sample with known high concentrations of the analyte. The operator performing the test will be blinded as to the identity of the samples (i.e. whether they are from cases of proven IA [samples containing the analyte] or controls [samples not containing the analyte]).

Twenty specimens representing a range of analyte concentrations (a minimum of 2, but preferably 5, analyte concentrations) and 20 to 50 specimens that lack the analyte will be tested.<sup>2,3,4</sup> Each specimen will be assayed in replicate as recommended by the third party.

1. Statistical Analyses

Results will be reported in a common 2 x 2 table format, Data Sheet 5 in Appendix A. Accuracy, sensitivity and specificity (for tests using animal or banked human samples), or positive and negative percent agreement (for tests using non-reference standards) and respective two-sided 95% score confidence intervals will be calculated per CLSI Document EP12-A and FDA Guidance Document *Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests*.<sup>4,5</sup>

The test will be considered verified if the sensitivity and specificity are within a certain percentage (as defined by the AsTeC Review Committee) below those claimed by the third party.

- K. After the completion of the replication study, a final GLP compliant report will be prepared and submitted to the Project Officer for review and approval.

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
Contract No. HHSN266200700023C  
Standard Operating Procedures

**REFERENCES:**

1. NCCLS. Evaluation of Precision Performance of Quantitative Measurement Methods: Approved Guideline – Second Edition. NCCLS document EP5-A2. 2004. NCCLS, Wayne, PA.
2. Elder, B. L., S. H. Hansen, J. A. Kellogg, F. J. Marsik, and R. J. Zabransky. 1997. Cumitech 31, Verification and validation of procedures in the clinical microbiology laboratory. Coordinating ed., B.W. McCurdy. ASM Press, Washington, D.C.
3. NCCLS. User Demonstration of Performance for Precision and Accuracy; Approved Guideline. NCCLS document EP15-A. 2001. NCCLS, Wayne, PA.
4. NCCLS. User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline. NCCLS document EP12-A. 2002. NCCLS, Wayne, PA.
5. FDA Guidance Document 1620 - *Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests*. March 13, 2007. U.S. Department of Health and Human Services Food and Drug Administration Center for Devices and Radiological Health.
6. FDA Guidance Document 1825 –*Class II Special Controls Guidance Document: Serological Assays for the Detection of Beta-Glucan*. September 23, 2004. U.S. Department of Health and Human Services Food and Drug Administration Center for Devices and Radiological Health.
7. NCCLS. Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach; Approved Guideline EP6-A. 2003. NCCLS, Wayne, PA.
8. NCCLS. Protocols for Determination of Limits of Detection and Limits of Quantitation; Approved Guideline – NCCLS document EP17-A2. 2004. NCCLS, Wayne, PA.

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
 Contract No. HHSN266200700023C  
 Standard Operating Procedures

**Appendix A. Sample Data Recording Sheets**

**Data Sheet #1: Precision Evaluation Experiment (a separate recording sheet will be needed for each analyte concentration tested)**

Concentration:

Operator:

Analyte:

Reagent Source/Lot:

Device:

Calibrator Source/Lot:

Third Party Claim of Repeatability Variance (SD):

Third Party Claim of Within Device Precision (SD):

Day	Date	Run 1			Run 2			Daily Mean
		Result 1	Result 2	Mean	Result 1	Result 2	Mean	
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
 Contract No. HHSN266200700023C  
 Standard Operating Procedures

**Data Sheet #2: Precision Experiment: Qualitative Tests**

Date of Testing: \_\_\_\_\_ Operator: \_\_\_\_\_  
 Sample/Matrix Type: \_\_\_\_\_  
 Analyte: \_\_\_\_\_ Reagent Source/Lot: \_\_\_\_\_  
 Third Party's Claimed Cutoff Concentration\*: \_\_\_\_\_  
 Device: \_\_\_\_\_ Calibrator Source/Lot: \_\_\_\_\_

Rank	Sample 1 (20% above cutoff)	Sample 2 (20% below cutoff)	Sample 3 (At cutoff*)
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
<b>% Positive Results</b>			
<b>% Negative Results</b>			

\* The concentration at which repeated tests on the same sample yield positive results 50% of the time and negative results for the other 50%.

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
 Contract No. HHSN266200700023C  
 Standard Operating Procedures

**Data Sheet #3: Limit of Detection Experiment (a separate recording sheet will be needed for each microbial variant tested)**

Date of Testing: \_\_\_\_\_ Operator: \_\_\_\_\_  
 Sample/Matrix Type: \_\_\_\_\_  
 Analyte: \_\_\_\_\_ Reagent Source/Lot: \_\_\_\_\_  
 LoD Sample Analyte Concentration: \_\_\_\_\_  
 Third Party's claimed LoB: \_\_\_\_\_  
 Third Party's claimed LoD: \_\_\_\_\_  
 Device: \_\_\_\_\_ Calibrator Source/Lot: \_\_\_\_\_

Rank	LoB Sample Value	LoD Sample Value
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
<b>Number of LoB samples exceeding claimed LoB</b>		
<b>Number (%) of LoD samples exceeding claimed LoB</b>		

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
Contract No. HHSN266200700023C  
Standard Operating Procedures

**Data Sheet #4: Linearity Evaluation**

Date of Testing:

Operator:

Concentration:

Sample/Matrix Type:

Analyte:

Reagent Source/Lot:

Device:

Calibrator Source/Lot:

<b>Dilution</b>	<b>Replicate #1</b>	<b>Replicate #2</b>	<b>Replicate #3</b>	<b>Replicate #4</b>	<b>Mean</b>
<b>1</b>					
<b>2</b>					
<b>3</b>					
<b>4</b>					
<b>5</b>					
<b>6</b>					
<b>7</b>					

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
 Contract No. HHSN266200700023C  
 Standard Operating Procedures

**Data Sheet #5: Accuracy Evaluation Experiment**

Date of Testing:

Operator:

Method: X

Sample/Matrix Type:

Analyte:

Reagent Source/Lot:

Device:

Calibrator Source/Lot:

		Known Sample Status (+/-)		Total
		Positive	Negative	
Method X Result	Positive	a*	b	a+b
	Negative	c	d	c+d
Total		a+c	b+d	n

\*The values that correspond to the letters in the table are used in the formulas outlined in NCCLS document EP12-A Evaluation of Qualitative Test Performance.

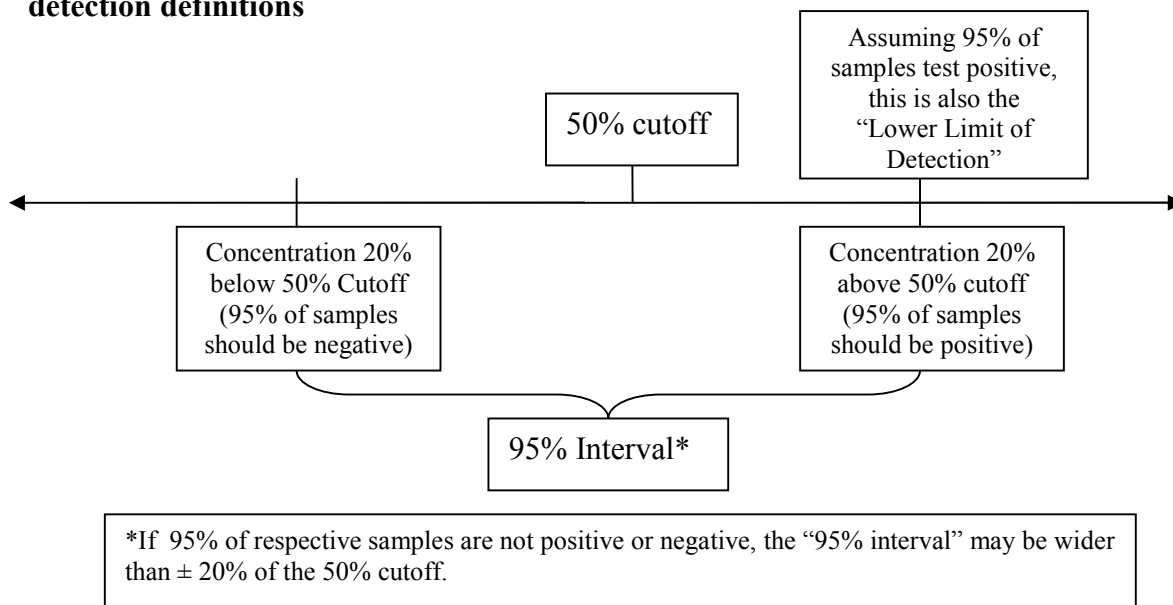
## Appendix B: Definitions

### 1. Determining Precision for Qualitative Assays

**50% Cutoff Point** - the analyte concentration at which repeated tests on the same sample yield positive results 50% of the time and negative results for the other 50%. The precision experiment described for qualitative assays in this document cannot define the 95% interval\* but rather are intended to determine if results that fall within +/- 20% of the 50% cutoff concentration are within the 95% interval for the test method. If the 95% interval does not fall within  $\pm 20\%$  of the 50% cutoff, additional testing is required to determine the actual 95% interval. This testing is performed in order to evaluate precision of a qualitative method *at analyte concentrations near the cutoff*.

**95% Interval for Cutoff Point** - The concentrations above and below the 50% cutoff point at which repeated results are 95% positive or 95% negative, respectively. This interval is sometimes referred to as the “indeterminate” range for qualitative tests.

**Figure 1. Relationship between cutoff points, intervals, and limit of detection definitions**



### 2. Determining/Verifying the Limit of Detection (LoD) and Limit of the Blank (LoB)

**Limit of Detection (LoD)**-the lowest amount of analyte in a sample that can be detected with (stated, in this case 95%) probability. This value can also be called the *lower limit of detection*.

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
Contract No. HHSN266200700023C  
Standard Operating Procedures

**Limit of the Blank (LoB)** - the highest measurement result that is likely to be observed (with a stated probability) for a blank sample.

Clinical Laboratory Diagnostics for Invasive Aspergillosis  
Contract No. HHSN266200700023C  
Standard Operating Procedures

ORIGINAL IMPLEMENTATION DATE: \_\_\_\_\_

APPROVED BY NIH NIAID Project Officer: \_\_\_\_\_ DATE \_\_\_\_\_

APPROVED BY PI/designee: \_\_\_\_\_ DATE \_\_\_\_\_

APPROVED BY Laboratory Coordinator: \_\_\_\_\_ DATE \_\_\_\_\_