COLLEGE of AMERICAN PATHOLOGISTS

Calibration Verification/ Linearity Program Users Guide

New for 2016

- **Two New CVL Programs**

  LN43 – Lamellar Body Count CVL

  LN44 – Fibrinogen CVL

- **Instrument Specific C-Reactive Protein Options**
  The LN12 C-Reactive Protein CVL survey is now an instrument-specific program with two options.

  LN12 – target range of 5-110 mg/L intended for Beckman Immage, Siemens Dimension, and Vitros instruments

  LN12E – target range of 5-320 mg/L intended for Abbott Architect, Beckman (except Immage), Roche, and Siemens (except Dimension) instruments

- **Expanded D-dimer Range**
  The LN42 D-dimer CVL survey has an expanded target range to improve analytical measurement range (AMR) coverage.

Table of Contents

How to Make the Most of Your User Guide 2

CVL Evaluations and Reports 3
- Understanding Your CVL Surveys Report Package 3
- Peer Group Summary 3
- Peer Performance Bar Charts 3
- Executive Summary 3
- Calibration Verification Evaluation 4
- Standard Linearity Evaluation 6
- Diluted/Extended Linearity Evaluation 8
- Linearity Troubleshooting Report 11

Summary of Regulatory Requirements 12

Calibration Verification Troubleshooting Guide 13

Linearity Troubleshooting Guide 15

Glossary of Terms 19

Frequently Asked Questions 21

Appendix 1: Worksheet for Calibration Verification Self-Evaluation 25

Appendix 2: CVL Investigation Checklist for Problematic Results 26

Appendix 3: Flow Charts 28
- Selection of Linearity Evaluation Type 28
- Standard Linearity Evaluation 29
- Diluted/Extended Linearity Evaluation When Results Include Extended Range (ER) Specimens 30
- Diluted/Extended Linearity Evaluation When Results Include Diluted Specimens 31

Appendix 4: Technical Guide to the CAP Linearity Evaluation 32

References 35

How to Make the Most of Your User Guide

The CVL Surveys User Guide includes several easy-to-use sections.

- **The CVL Report Package (pages 3–11)—**Sample report pages and a summary of the calibration verification and linearity evaluations. This section of the User Guide highlights key concepts and terminology.

- **Regulatory Requirements (page 12)—**Information about regulatory requirements pertaining to the CVL Surveys.

- **Calibration Verification and Linearity Troubleshooting Guides (pages 13–18)—**A collection of examples are provided if you have any problems with your evaluations. You can identify the example report that is most similar to your results and that classifies your issue and suggests what actions should be taken.

- **Reference Information: Glossary and FAQ (pages 19–24)—**A complete glossary of calibration verification and linearity terms and a listing of frequently asked questions.

- **Additional Information: Appendices (pages 25–34)—** More detailed explanations of the evaluation protocols.

  - Appendix 1 provides a self-evaluation worksheet for calibration verification.
  - Appendix 2 contains a checklist for investigating common calibration verification and linearity problems.
  - Appendix 3 features detailed flow charts outlining the logic and procedural steps in the linearity evaluations.
  - Appendix 4 offers a more advanced discussion of the statistical methods.

Thank you for participating in the CAP CVL program!

CVL Evaluations and Reports

Understanding Your CVL Surveys Report Package

Your CVL report package includes the following individual and peer group performance summaries:
- Peer Group Summary
- Peer Performance Bar Charts
- Executive Summary
- Calibration Verification Evaluation
- Linearity Evaluation — Standard or Diluted/Extended
- Linearity Troubleshooting Report (if applicable)

This section provides a detailed description of each component.

Peer Group Summary

The Participant Summary contains a listing of peer means and coefficients of variation (CVs) for all specimens. This information can be useful for troubleshooting problems with your calibration verification and linearity results. We exclude outliers when we calculate peer means and CVs. Elevated CVs can indicate problems in your peer group related to specimen mishandling, elevated interlaboratory variability, or possible bimodality. Beginning in 2015, those CVL Surveys with SI processing will summarize participant data in both conventional and SI units.

Peer Performance Bar Charts

The Participant Summary contains bar charts summarizing calibration verification and linearity performance by peer group. The peer performance bar charts can also be useful to identify possible peer-level problems. Review these charts for information on overall performance as well as performance within your peer group.

Executive Summary

The Executive Summary lists your calibration verification and linearity results for all analytes in a given Survey. The Executive Summary also indicates if your verified and/or linear ranges do not include all of your reported specimen results. If we have omitted results for low and/or high specimen levels to obtain a verified or linear evaluation, it is important to review your evaluation report in more detail to determine if excluded specimens reveal possible analytical problems. In addition, the Executive Summary will indicate if your verified and/or linear ranges include any diluted specimens. Finally, the Executive Summary includes a note if we are unable to complete your calibration verification or linearity evaluation due to lack of a peer group (No Peer Group) or insufficient data (Insufficient Data for Evaluation).

Calibration Verification Evaluation

The calibration verification evaluation compares participant results to method-appropriate target values. Most commonly, target values equal the peer means calculated from participants in the current mailing. In accuracy-based surveys, we assign target values by reference or definitive method.

EVALUATION
ORIGINAL

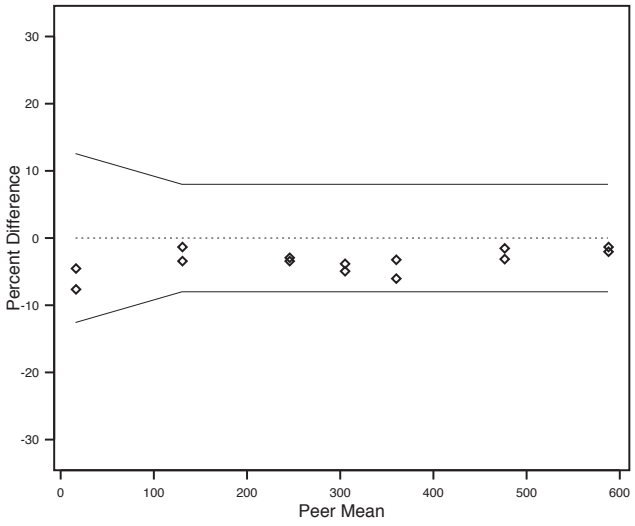
LN11-B Serum Ethanol Calibration Verification/Linearity
Serum Ethanol mg/dL Calibration Verification Evaluation

1 **Evaluation Result: Verified from 14.95 to 577.65**
Peer Method: GAS CHROMATOGRAPHY

Allowable Error: 8% or 2 mg/dL, whichever is greater

Specimen	Assay 1	Assay 2	Your Mean	Peer Mean	2 Peer N	3 Difference	4 Allowable Error
LN11-08	14.7	15.2	14.95	15.91	12	-0.96 mg/dL	±2 mg/dL
LN11-09	128.5	125.7	127.10	130.14	12	-2.3%	± 8.0%
LN11-10	236.8	238.1	237.45	245.23	12	-3.2%	± 8.0%
LN11-11	289.7	293.0	291.35	304.65	12	-4.4%	± 8.0%
LN11-12	338.1	348.2	343.15	359.63	12	-4.6%	± 8.0%
LN11-13	461.1	468.9	465.00	475.95	12	-2.3%	± 8.0%
LN11-14	579.5	575.8	577.65	587.39	11	-1.7%	± 8.0%

5 Calibration Verification Plot: Percent Differences with Allowable Error Limits



6 **Peer Results Summary Table**

7 Peer Group Size: 12

Range	Calibration Verification		Linearity Evaluation		
	% Verified	% Different	% Linear	% Nonlinear	% Imprecise
LN11-08 - 14	83.3	8.3	91.7	0.0	0.0
LN11-08 - 13	8.3	0.0	8.3	0.0	0.0

Your calibration verification evaluation result can be Verified or Different. Your verified range may not include all reported specimens.

- **Verified** means all differences between your mean and the peer mean are within the calibration verification allowable error limits in the range specified.

- **Different** means the difference between your mean and the peer mean exceeds the calibration verification allowable error limit for at least one specimen.

If the calibration verification evaluation is Different over the full range of submitted results, we run subsequent evaluations with results at the high and low ends systematically omitted. At the low end, we only remove the results for the lowest specimen. We provide a final evaluation for the first range of specimens that is Verified and includes the required number of consecutive data sets stated in the kit instructions. If there are no verified ranges, the final evaluation result is Different. Even if your calibration verification evaluation result is Verified in a range that includes all of your reported specimens, we recommend reviewing your differences from your peer means and your calibration verification plot. This information can provide early insight into problems with calibration and repeatability.

Key to Calibration Verification Report

1. Your evaluation result will be either Verified or Different. If your evaluation result is Verified, we list your means of the low and high specimens included in the verified range.
2. When we use peer means for target values, we list the number of participants reporting results for each specimen. These numbers provide additional information about your peer group performance.
3. The difference is 100 x (Your Mean – Peer Mean) / (Peer Mean). If the alternative error limit in absolute units is greater than the error percentage, then the difference will be expressed in absolute units.
4. The allowable error is the larger of the two limits listed above the table for that particular specimen.
5. Your calibration verification plot provides a visual comparison between your results and peer means. The appropriate error limits are also presented for comparison. See the Calibration Verification Troubleshooting Guide section of the User Guide for assistance in interpreting common patterns.
6. The peer results summary provides information about the performance of others in your peer group. Listed are the final evaluation ranges along with the percent of peer group members for each range. We calculate the percentages using the peer group size as the denominator.
7. The peer group size is the highest count of reported results for the Peer N's reported by specimen in the top table. We use the peer group size for the denominator when we calculate the percentages in the peer results summary table.

Standard Linearity Evaluation

In the linearity evaluation, we plot the relative concentrations on the horizontal axis and the measured results on the vertical axis. We fit a straight line through the points and determine if this line adequately describes the relationship between your reported results and the relative concentrations.

Your linearity evaluation result can be Linear, Nonlinear, or Imprecise (Poor Repeatability and/or Fit). Because we may evaluate smaller ranges in your data, your linear range may not include all reported specimens.

EVALUATION
ORIGINAL

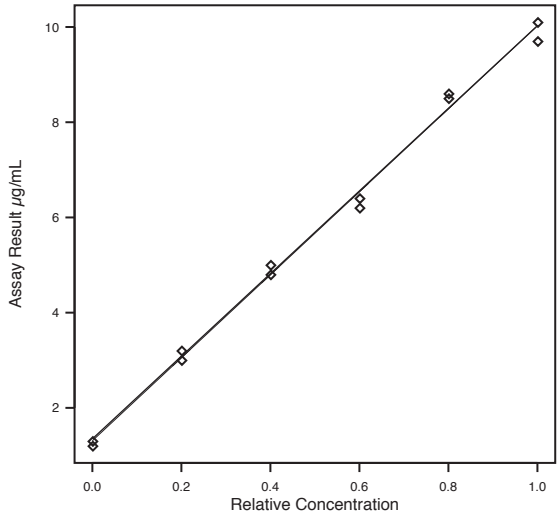
LN3-A TDM Calibration Verification/Linearity
Gentamicin µg/mL Linearity Evaluation

1 **Evaluation Result: Linear from 1.25 to 9.90**
Instrument: SIEMENS ADVIA CHEM SYST
Method: SYVA EMIT 2000

2 Evaluation Type: Standard
3 Goal for Total Error (TE): 20%
4 Mean of Included Results: 5.67 µg/mL

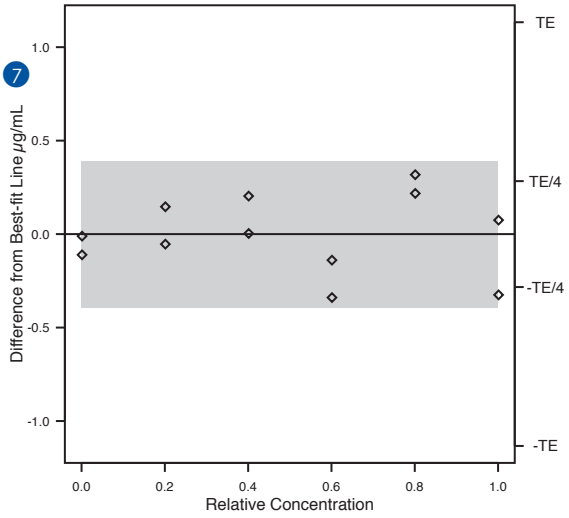
Specimen	Assay 1	Assay 2	Your Mean	5 Best-fit Target	6 Relative Concentration
LN3-01	1.3	1.2	1.25	1.31	0.000
LN3-02	3.2	3.0	3.10	3.05	0.200
LN3-03	4.8	5.0	4.90	4.80	0.400
LN3-04	6.2	6.4	6.30	6.54	0.600
LN3-05	8.5	8.6	8.55	8.28	0.800
LN3-06	9.7	10.1	9.90	10.02	1.000

Linearity Plot 1: Reported Results with Best-fit Line



Linearity Plot 2: Differences with Limits of Acceptable Imprecision

8



9
◇ Included in best-fit line
○ Excluded from best-fit line

Your evaluation result can be Linear, Nonlinear of Imprecise. Your linear range may not include all reported specimens.

- **Linear** means your reported results meet the criteria for acceptable linearity in the specified range. While there may be evidence of small deviations from linearity, results are within acceptable limits for nonlinearity and imprecision.
- **Nonlinear** means your reported results do not meet the criteria for acceptable linearity in any of the ranges evaluated within the algorithm. The selection of ranges depends on whether your results include extended range or diluted specimens.
- **Imprecise (Poor Repeatability and/or Fit)** means your results display too much variability to permit a reliable determination of linearity. This evaluation result reflects either large differences between specimen replicates (poor repeatability) or large differences between specimen means and the best-fit line or curve (poor fit).

In the standard evaluation, if the evaluation is Nonlinear or Imprecise (Poor Repeatability and/or Fit) over the full range of submitted results, we run subsequent evaluations with results at the high end systematically omitted. A final evaluation result is given for the first range of specimens that is linear and includes the required number of consecutive data sets stated in the kit instructions. If there are no linear ranges, we base the evaluation on the full range of submitted results, and it will be either Nonlinear or Imprecise (Poor Repeatability and/or Fit).

If you receive a Nonlinear or Imprecise (Poor Repeatability and/or Fit) evaluation, your report package will include a linearity troubleshooting report. The troubleshooting report includes an interpretation and set of suggested actions based on your results.

Key to Standard Linearity Evaluation

1. Your evaluation result will be Linear, Nonlinear, or Imprecise (Poor Repeatability and/or Fit). If your evaluation result is Linear, we list your means of the low and high specimens included in the linear range.
2. There are two linearity evaluation types: standard and diluted/extended.
3. We use the goal for total error to specify limits on both acceptable imprecision and nonlinearity. We estimate both imprecision and nonlinearity using averages taken across all specimen levels. The limits for both are slightly larger than one-fourth the goal for total error evaluated at the mean of the assay results included in your best-fit line. The maximum total error goal used for the linearity evaluation is 25%.
4. We provide the mean of included results to assist with the interpretation of your linearity results. We calculate the limits of both acceptable imprecision and nonlinearity using the goal for total error and the mean of included results.
5. The best-fit target is the fitted value from the best-fit line shown in Linearity Plot 1 for each specimen.
6. The relative concentration is the proportional concentration of analyte in the specimen. In most Surveys, we standardize the proportional concentration to zero for the lowest specimen and one for the highest specimen.
7. These plots use the units of measure from your test result forms. In some cases, the unit of measure is a percent.
8. The labels on this axis show the differences on the scale of total error units. One total error unit is equal to (Goal for Total Error / 100) x (Mean of Included Results).
9. In the standard evaluation, results are excluded from the best-fit line if the evaluation result is not linear in the full range of reported results.

Diluted/Extended Linearity Evaluation

You may receive a diluted/extended linearity evaluation if your results include at least one extended range or diluted specimen. An extended range specimen follows a large gap in the relative concentrations; the diluted/extended evaluation prevents these results from having a large influence on the estimated best-fit line. You will receive a diluted/extended evaluation for a diluted specimen when two conditions are met: 1) You indicate that you diluted both assays for the specimen, and 2) You have an adequate number of undiluted specimen results that can be evaluated using the standard linearity evaluation. Generally, this requirement is four sets of specimen results.

In a diluted/extended evaluation, we first evaluate the results excluding the diluted or extended range specimens using the standard linearity protocol. If this initial evaluation is linear, we then evaluate the first diluted or extended range specimen for consistency with the fitted line. The evaluation algorithm calculates the predicted value for the diluted or extended range specimen by extending the line fit in the first step to the additional specimen. We compare the predicted value to the mean of your diluted or extended range specimen results. If the difference is smaller than the allowable error limit, we will extend the linear range to the diluted or extended range specimen.

The diluted/extended evaluation also includes a screen for large replicate differences for the diluted or extended range results. If the difference between replicates for a diluted or extended range specimen exceeds the replicate limit, the linear range will not be extended, regardless of the agreement between your mean and the predicted value.

If the evaluation is Nonlinear or Imprecise (Poor Repeatability and/or Fit) over the full range of submitted results excluding the diluted or extended range specimens, we run subsequent evaluations with results at the high end systematically omitted. We determine an intermediate evaluation result for the first solution range that is linear and includes the required number of consecutive data sets stated in the kit instructions. If none of the ranges evaluated are linear, the evaluation given is based on the full range of regular specimens and will be either Nonlinear or Imprecise (Poor Repeatability and/or Fit). Starting in 2015, if the linearity evaluation result includes diluted specimens, the result will be displayed on two lines to differentiate the standard evaluation results from the diluted/extended range results.

NOTE: See the Diluted/Extended Linearity Evaluation flow chart in Appendix 3 for more information.

Key to Differences Between Diluted/Extended and Standard Linearity Evaluations When Results Include Extended Range Specimens

EVALUATION
ORIGINAL

LN13-B Blood Gas Calibration Verification/Linearity
PO2 mm Hg Linearity Evaluation

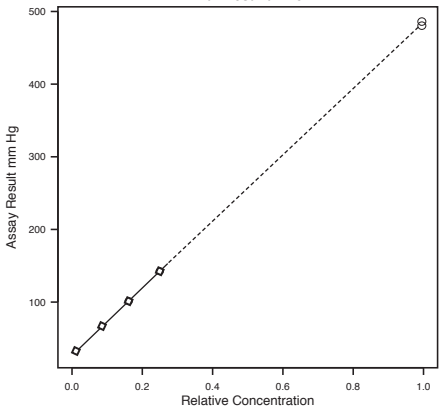
1 Evaluation Result: Linear from 29.0 to 485.5
Instrument: SIEMENS RAPIDLAB 1260
Barometric Pressure: > OR = 720 mm Hg

Evaluation Type: Diluted/Extended
Goal for Total Error (TE): 16%
Mean of Included Results: 83.1 mm Hg

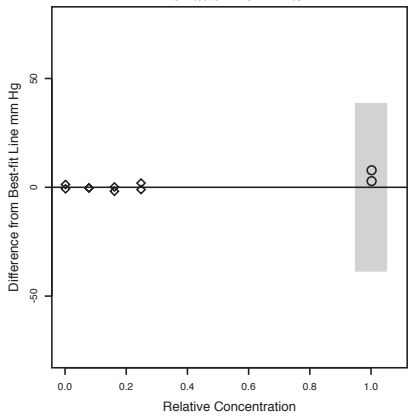
Specimen	Assay 1	Assay 2	Your Mean	3 Best-fit Target	Relative Concentration
LN13-06	28	30	29.0	28.6	0.000
LN13-07	63	63	63.0	63.2	0.077
LN13-08	101	99	100.0	100.7	0.160
LN13-09	142	139	140.5	140.0	0.247
§ LN13-10	483	488	485.5	480.0	1.000

2 § Extended range specimen

Linearity Plot 1: Reported Results with Best-fit Line



Linearity Plot 2: Differences with Allowable Error Limits



Best-fit line
Extended line
Included in best-fit line
Excluded from best-fit line

4 Your Extended Range Specimen(s) Analysis

Specimen	Replicate Difference [1]	Your Mean	Best-fit Target	Difference From Target	Allowable Error [2]	Evaluation
LN13-10	5	485.5	480.0	5.5	± 38.4	Included in linear range

[1] Limit for replicate difference = 5.3% of Best-fit Target
[2] Allowable Error = 8% of Best-fit Target

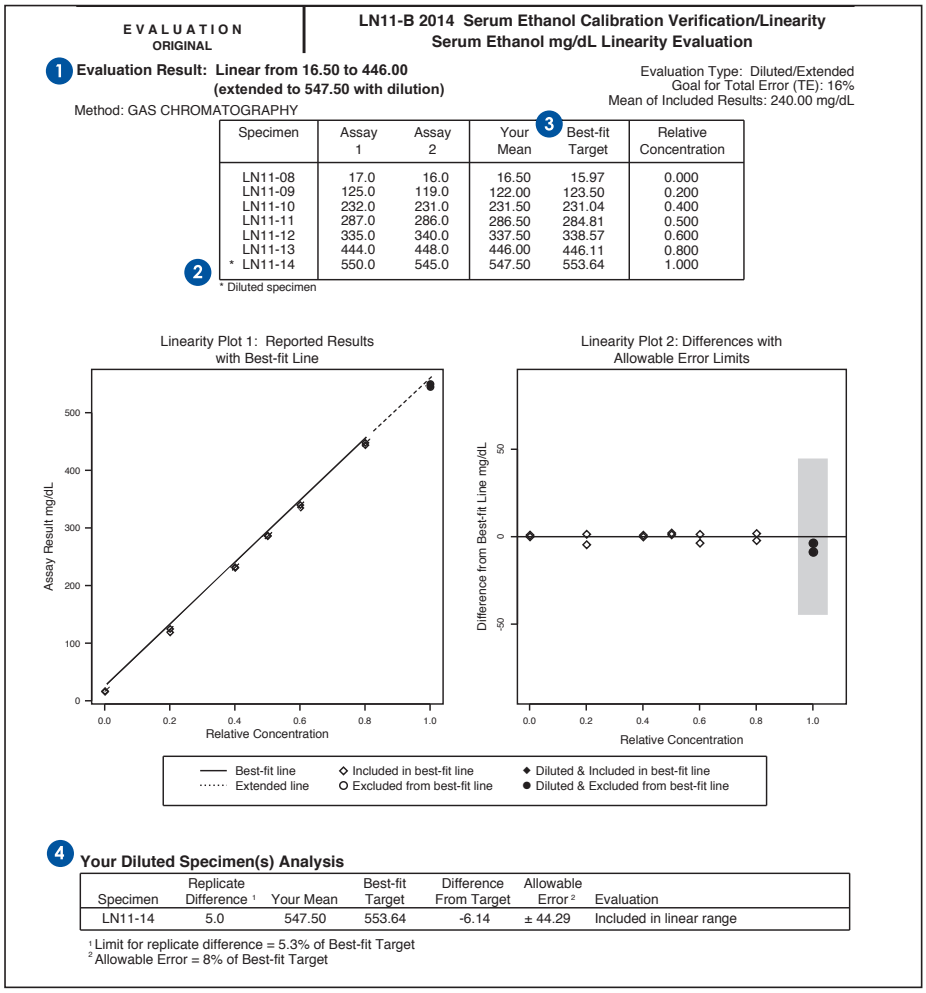
1 Your linear range may contain the extended range specimen(s). No range is given if your result is Nonlinear or Imprecise (Poor Repeatability and/or Fit).

2 Extended range specimens are identified in the Specimen column.

3 The best-fit target is the fitted value from the best-fit line shown in Linearity Plot 1. We do not include extended range specimen results when finding the best-fit line. We calculate the best-fit targets for the extended range specimens from the line fit to the nonextended range specimens.

4 If your nonextended range results are linear in the full range, we evaluate your extended range specimens for consistency with the best-fit line. The linear fit may be extended to include extended range specimens that are within the allowable error limits. If the difference between replicates for the extended range specimen exceeds the replicate limit, the linear range will not be extended. If your evaluation is not linear, we will not evaluate your extended range specimen results for linearity and your report will not include an extended range specimen analysis. See Section 5 of Appendix 4 for information on the determination of the allowable error limits.

Key to Differences Between Diluted/Extended and Standard Linearity Evaluations When Results Include Diluted Specimens



EVALUATION
ORIGINAL

LN11-B 2014 Serum Ethanol Calibration Verification/Linearity
Serum Ethanol mg/dL Linearity Evaluation

1 **Evaluation Result: Linear from 16.50 to 446.00**
(extended to 547.50 with dilution)

Evaluation Type: Diluted/Extended
Goal for Total Error (TE): 16%
Mean of Included Results: 240.00 mg/dL

Method: GAS CHROMATOGRAPHY

Specimen	Assay 1	Assay 2	Your Mean	3 Best-fit Target	Relative Concentration
LN11-08	17.0	16.0	16.50	15.97	0.000
LN11-09	125.0	119.0	122.00	123.50	0.200
LN11-10	232.0	231.0	231.50	231.04	0.400
LN11-11	287.0	286.0	286.50	284.81	0.500
LN11-12	335.0	340.0	337.50	338.57	0.600
LN11-13	444.0	448.0	446.00	446.11	0.800
* LN11-14	550.0	545.0	547.50	553.64	1.000

2 * Diluted specimen

Linearity Plot 1: Reported Results with Best-fit Line

Assay Result mg/dL

Relative Concentration

Linearity Plot 2: Differences with Allowable Error Limits

Difference from Best-fit Line mg/dL

Relative Concentration

Best-fit line
Extended line
Included in best-fit line
Excluded from best-fit line
Diluted & Included in best-fit line
Diluted & Excluded from best-fit line

4 **Your Diluted Specimen(s) Analysis**

Specimen	Replicate Difference [1]	Your Mean	Best-fit Target	Difference From Target	Allowable Error [2]	Evaluation
LN11-14	5.0	547.50	553.64	-6.14	± 44.29	Included in linear range

[1] Limit for replicate difference = 5.3% of Best-fit Target
[2] Allowable Error = 8% of Best-fit Target

1 Your linear range may contain your diluted specimen(s). We do not specify a range if your result is Nonlinear or Imprecise (Poor Repeatability and/or Fit).

2 Diluted specimens are identified by an asterisk in the Specimen column. Both assays must be diluted for the specimen to be designated as a diluted specimen.

3 The best-fit target is the fitted value from the best-fit line shown in Linearity Plot 1. We do not include diluted specimen results when finding the best-fit line. We calculate the best-fit targets for the diluted specimens from the line fit to your undiluted results.

4 If your undiluted results are linear in the full range, we evaluate your diluted specimens for consistency with the best-fit line. If your diluted specimens are within the allowable error limits, we extend your linear range to contain consecutive diluted specimens. If the differences between replicates for any diluted specimen exceeds the replicate limit, the linear range will not be extended. If your evaluation is not linear for the full range of undiluted specimens, we will not evaluate your diluted specimens for linearity and your report will not include a diluted specimen analysis. See Section 5 of Appendix 4 for information on the determination of the allowable error limits.

Linearity Troubleshooting Report

You will receive a linearity troubleshooting report for each analyte with a linearity evaluation result of Nonlinear or Imprecise (Poor Repeatability and/or Fit).

EVALUATION
ORIGINAL

LN2-A Chemistry/Lipid/Enzyme Calibration Verification/Linearity
CO2 mmol/L Linearity Troubleshooting Report

Evaluation Result: Imprecise (Poor Repeatability and/or Fit)

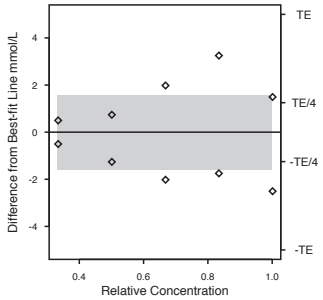
1 **Interpretation:**
Your results have large differences between assay replicates.

2 **Suggested Actions:**
Review specimen handling procedures. Incomplete thawing, inadequate mixing, or improper temperatures can cause problems.

Check for evidence of test system malfunction. Pipetting errors, temperature fluctuations, or optical system problems can contribute to large analytical errors.

If unable to resolve, contact your method manufacturer.

3 Troubleshooting Report Plot: Differences with Allowable Error Limits



Document investigation and corrective action here or on the CVL Investigation Checklist:

Reviewed by: Date:

Key to Linearity Troubleshooting Report

1 The interpretation describes the pattern of deviations from linearity in your data. Visual inspection of your difference plot should confirm the interpretation.

2 Suggested actions are provided based on the interpretation of your deviations from linearity.

3 The troubleshooting report plot shows the deviations from the best-fit line. This plot may not show your diluted results since the interpretation and suggested actions focus on problems within your analytical measurement range (AMR). If you have fewer undiluted specimens than required for a linearity evaluation, your difference plot will include both undiluted and diluted results.

Summary of Regulatory Requirements

The CAP Calibration Verification/Linearity Surveys provide specimens and statistical evaluations of the reported results for verification of your current calibration settings as well as for assessing the analytical measurement range (AMR) of your laboratory method. The CVL Surveys satisfy the requirements for scheduled calibration verification and AMR verification as specified in the CAP Laboratory Accreditation Program (LAP) and Current CLIA Regulations Section 493.1255 for most analytes.[1,2] As defined in the CAP LAP Chemistry and Toxicology Checklist, the AMR is the range of analyte values that a method can directly measure on the specimen without any dilution, concentration, or other pretreatment that is not part of the usual assay process.

The CAP CVL Surveys materials are intended to cover broad ranges of concentrations or activities to challenge assays from near the low end to the top of the AMR for most analytes and methods. To fulfill CAP LAP requirements for AMR verification, you must know your analytical measurement range for each analyte, and CVL Survey concentrations must fall near the low, middle, and high values of your AMR. If specimens do not meet these requirements, you may find additional information on regulatory requirements in the Frequently Asked Questions section of this document.

You may use either your calibration verification or linearity evaluation, or a combination of both, to fulfill requirements for calibration verification or AMR verification. You must confirm that the performance limits specified in the evaluation are acceptable to your laboratory. For AMR verification, you must confirm that the evaluation includes specimens near the low, midpoint, and high values of the AMR. If your CVL Surveys evaluation does not satisfy your own criteria, you have the option of performing a self–evaluation with alternate target values and/or with modified limits of allowable error.

NOTE: See the calibration verification and linearity troubleshooting guides for suggested actions if you have problematic results. Appendix 1 provides a worksheet for calibration verification self-evaluation. Appendix 2 provides a detailed investigation checklist.

Coagulation

According to the CAP LAP Hematology and Coagulation Checklist, coagulation tests based on direct measurement of an analyte require AMR verification at least every six months.[8] Specifically, calibrated tests that directly measure activity or concentration of an analyte by enzyme immunoassay, immunoturbidity, or chromogenic methods require AMR verification. Clot-based tests do not require AMR verification.

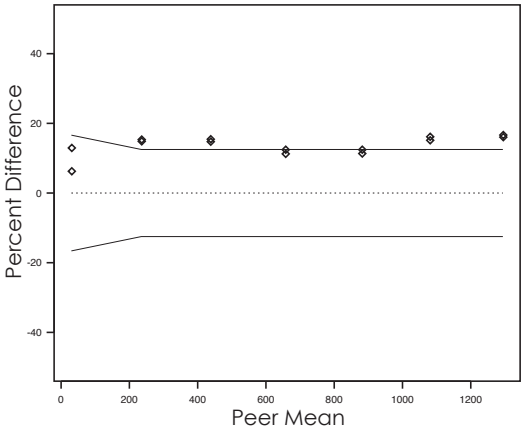
Hematology and Flow Cytometry

According to the CAP LAP Hematology and Coagulation Checklist, linearity studies are not required to satisfy calibration verification for complete blood count (CBC) instruments. Similarly, the Flow Cytometry Checklist does not contain explicit requirements for confirmation of linearity for enumeration of blood lymphocytes. In an effort to standardize reports and documentation, we use the CAP LAP Chemistry and Toxicology Checklist terminology for most CVL program documents. Participants should be aware that specific accreditation and regulatory requirements could vary.

Calibration Verification Troubleshooting Guide

This troubleshooting guide provides suggested actions if you receive a calibration verification evaluation result of Different, or if your evaluation result is Verified over a range that does not include all of your reported results. To use this guide, determine which of the following examples is most similar to your calibration verification plot. Refer to the corresponding suggested actions, in conjunction with the CVL Investigation Checklist for Problematic Results, to investigate possible causes and corrective actions.

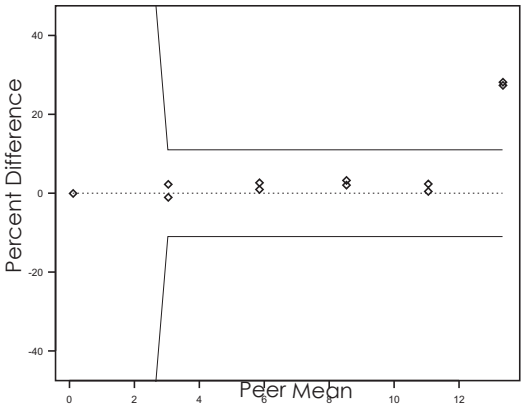
Constant Bias



Review the **ANALYTICAL** section of the investigation checklist. Analytical problems that produce a constant bias may be due to a calibration error. Recalibration may be needed.

Review the **CLERICAL** section of the investigation checklist. Clerical errors that result in a constant bias are likely due to units of measure or decimal place errors, or incorrect peer group assignment. Clerical errors may indicate a need for additional staff training.

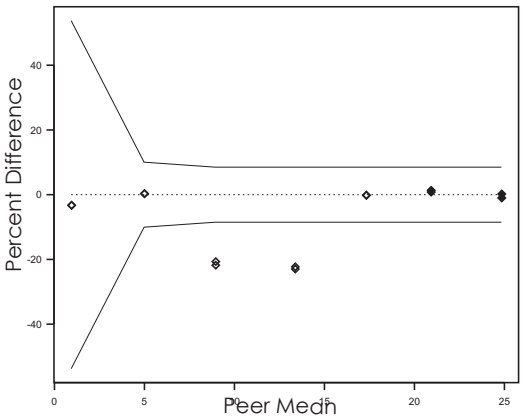
Problems With Low or High Specimens



Review the **ANALYTICAL** section of the investigation checklist. Analytical problems that appear at the low or high end may be due to recovery issues near or at the analytical measurement range (AMR) limits. If your results were diluted, review your laboratory's dilution protocol.

Review the **SPECIMEN HANDLING** section of the investigation checklist. Problems at the low or high end may also indicate sample degradation.

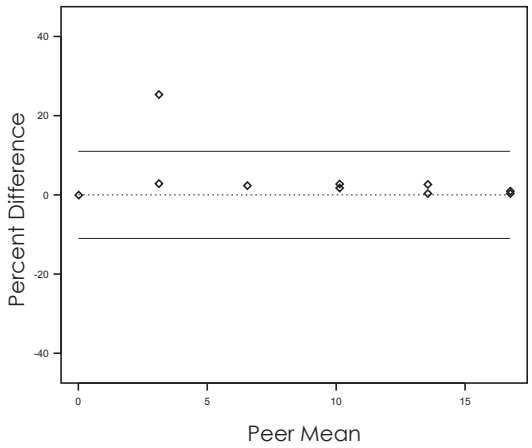
Problems With Middle Specimens



Review the **SPECIMEN HANDLING** section of the investigation checklist. Specimen handling problems may be due to tests performed on an incorrect vial, mixing or reconstitution problems, or improper specimen storage. Check whether special instructions were performed correctly.

Review the **CLERICAL** section of the investigation checklist. Clerical errors that show inconsistent recoveries for the middle specimens are likely due to a fax scanning, transcription, or result entry errors. If you suspect a fax scanning error, you must contact the CAP Customer Contact Center. If you suspect a transcription or result entry error, consider additional staff training.

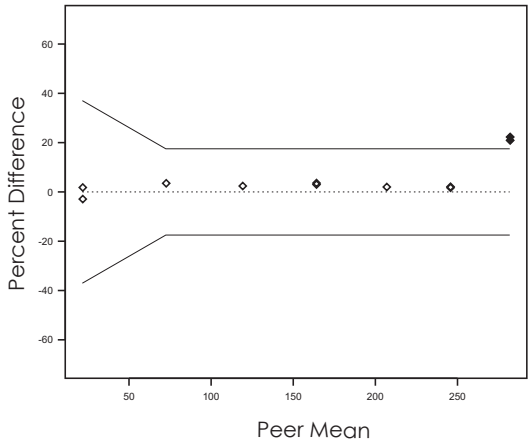
Large Difference Between Replicates for a Single Specimen



Review the **CLERICAL** section of the investigation checklist. Clerical errors that may cause a large difference between replicates for a single specimen are likely due to fax scanning, transcription, or result entry errors. If you suspect a fax scanning error, you must contact the CAP Customer Contact Center. If you suspect a transcription or result entry error, additional staff training may be needed.

Review the **SPECIMEN HANDLING** section of the investigation checklist. Specimen handling problems may be due to tests performed on an incorrect vial, mixing, or improper specimen storage. Check whether special instructions were performed correctly.

Problems With Diluted Specimens

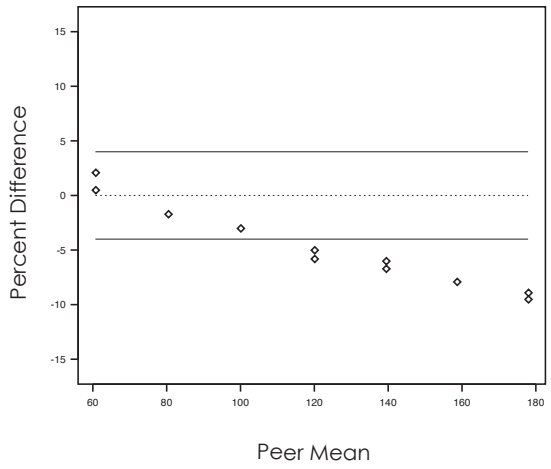


Review the **ANALYTICAL** section of the investigation checklist. Check whether the dilution protocol was followed (eg, dilution factor, diluents used). Confirm that the autodiluter is functioning correctly. If you suspect that the autodiluter is not functioning properly, you must contact the instrument manufacturer.

Review the **CLERICAL** section of the investigation checklist. Clerical errors with diluted specimens are likely due to use of an incorrect dilution factor.

Dilution errors resulting from a failure to adhere to protocol may indicate a need for additional staff training.

Proportional Bias



Review the **ANALYTICAL** section of the investigation checklist. Analytical problems that demonstrate increasing or decreasing bias may be due to a calibration error. Recalibration may be needed.

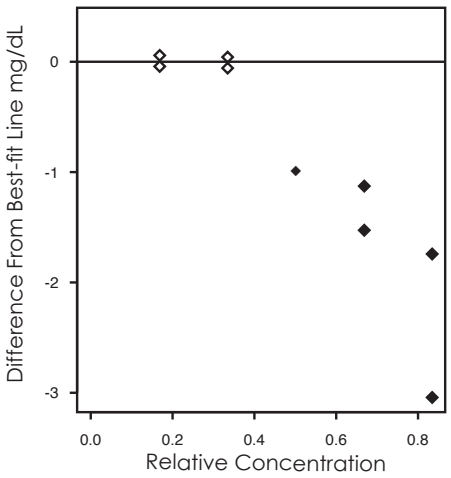
Review the **SPECIMEN HANDLING** section of the investigation checklist. Specimen handling problems that result in proportional bias may be due to mixing problems or improper storage.

Linearity Troubleshooting Guide

If you receive a linearity evaluation result of Nonlinear or Imprecise (Poor Repeatability and/or Fit), your report package will include a Linearity Troubleshooting Report. This troubleshooting report provides an interpretation and set of suggested actions based on the deviations from linearity displayed in your results. The following examples show diagnostic plots, interpretations, and suggested actions taken from recent linearity troubleshooting reports.

You may receive an evaluation result of Linear over a range that does not include all of your reported results. This outcome will be noted on your Executive Summary, but you will not receive a troubleshooting report. The final example in this guide shows a linear evaluation result over a partial range of specimens with a set of suggested actions. We recommend that you review these suggested actions when your linear range does not include all of your reported specimens.

Problems With Diluted Results



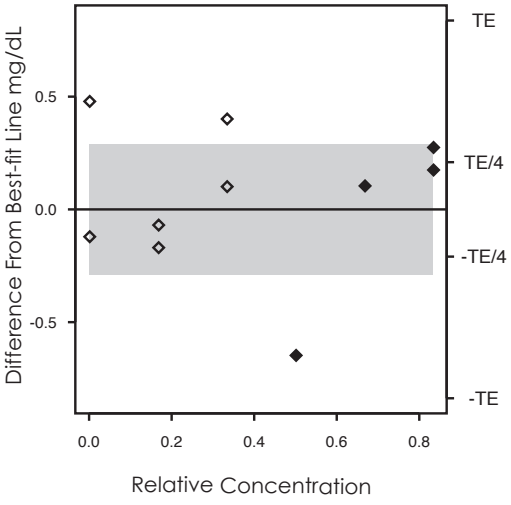
Review your dilution protocol, including appropriateness of the diluents.

Ensure that the dilution protocol is followed.

Rely on calibration verification results for undiluted specimens.

If your calibration verification evaluation result is Different, recalibrate and consider running a new linearity study.

Indeterminate—Problems Within AMR and/or With Diluted Results

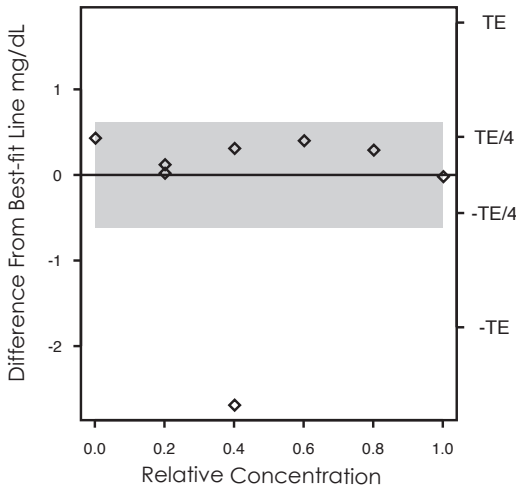


Review your dilution protocol and ensure the protocol is being followed.

Rely on calibration verification results for undiluted specimens.

If your calibration verification evaluation result is Different, recalibrate and consider rerunning a linearity study.

Large Difference Between Replicates for a Single Specimen

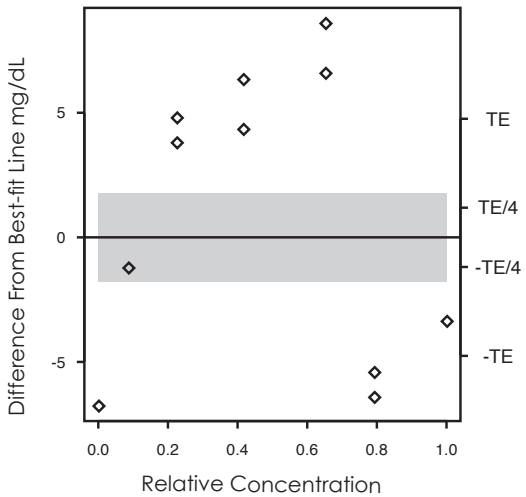


Rule out clerical error.

Review specimen handling procedures. Incomplete thawing, inadequate mixing, or improper temperatures can cause problems.

Review your testing procedure to ensure that the correct sample was analyzed and reported.

Reported Results Not in Sequential Order



Rule out clerical error, including the possibility that your samples were reported in the wrong order.

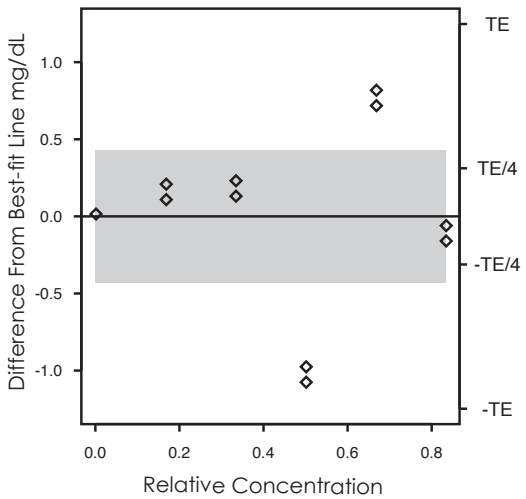
Check the AMR. Verify that you are not reporting results beyond your upper or lower AMR limit.

- Loss of recovery may be seen when values are close to the AMR limits.
- Check for dilution errors if values outside the AMR were diluted.

Review specimen handling procedures. Incomplete thawing, inadequate mixing, or improper temperatures can cause problems.

Review your testing procedure to ensure that the correct sample was analyzed and reported.

Poor Fit Between Results and Best-fit Line or Curve



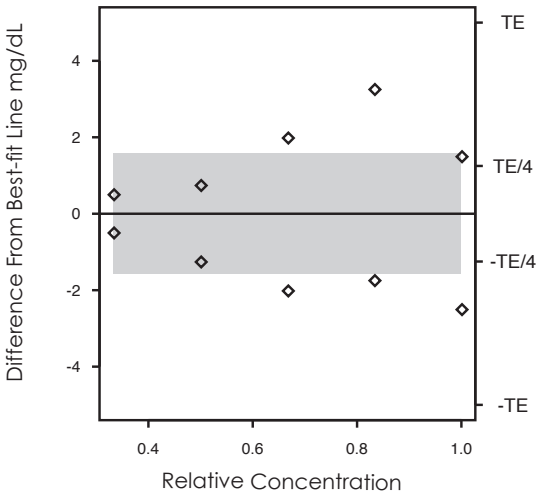
Check if both assays are equal. You may need to update your testing procedures—both assays from each vial should be tested within the same run.

Review specimen handling procedures. Incomplete thawing, inadequate mixing, or improper temperatures can cause problems.

Review your testing procedure to ensure that the correct sample was analyzed and reported.

If unable to resolve, consider recalibration.

Large Differences Between Replicates



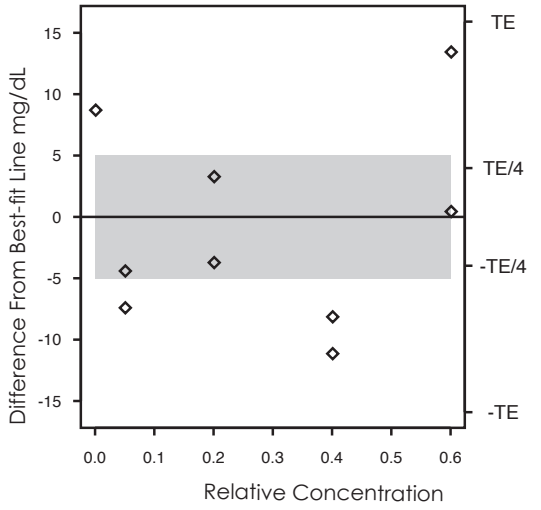
Check if one set of assays is consistently greater than the other. You may need to update your testing procedures—both assays from each vial should be tested within the same run.

Review specimen handling procedures. Incomplete thawing, inadequate mixing, or improper temperatures can cause problems.

Check for evidence of test system malfunction. Pipetting errors, temperature fluctuations, or optical system problems can contribute to large analytical errors.

If unable to resolve, consider recalibration.

Poor Agreement With the Best-fit Line or Curve and Large Differences Between Assay Replicates

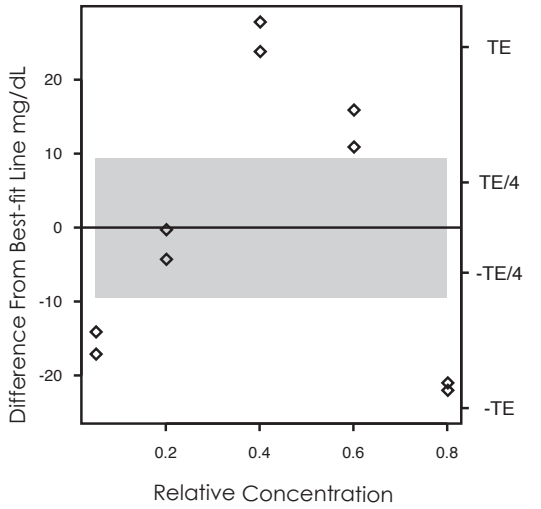


Review specimen handling procedures. Incomplete thawing, inadequate mixing, or improper temperatures can cause problems.

Check for evidence of test system malfunction. Pipetting errors, temperature fluctuations, or optical system problems can contribute to large analytical errors.

If unable to resolve, consider recalibration.

Nonlinear Results



Review specimen handling procedures. Incomplete thawing, inadequate mixing, or improper temperatures can cause problems.

Check the AMR. Verify that you are not reporting results beyond your upper or lower AMR limit.

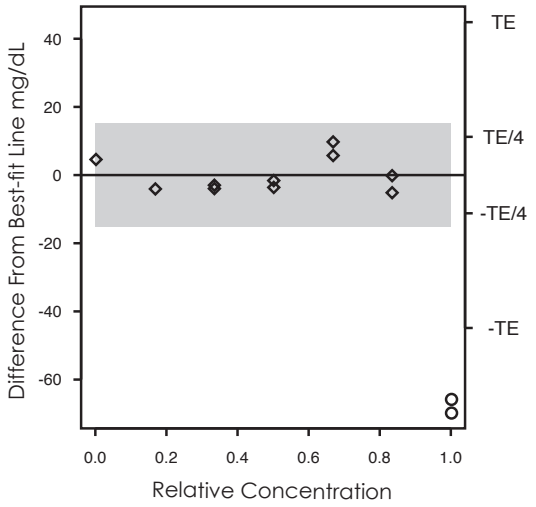
- Loss of recovery may be seen when values are close to the AMR limits.
- Check for dilution errors if values outside the AMR were diluted.

Check your peer group performance to see if there is nonlinear recovery for your peer group.

Check for evidence of test system malfunction.

- Aging reagents may cause low or high recovery at any level.
- Check calibration. Recalibration may be needed.
- Recalibration with a different lot may be needed.
- After recalibration, consider running a new linearity study to confirm that the problems have been resolved.

Linear Range Does Not Include All Reported Specimens



Review specimen handling procedures. Incomplete thawing, inadequate mixing, or improper temperatures can cause problems.

Check the AMR. Verify that you are not reporting results beyond your upper or lower AMR limit.

- Loss of recovery may be seen when values are close to the AMR limits.
- Check for dilution errors if values outside the AMR were diluted.

Check your Peer Results Summary Table to see if there are similar results for others in your peer group.

Check for evidence of test system malfunction. Aging reagents may cause low or high recovery at any level.

Glossary of Terms

Acceptable Imprecision Limits for Linearity Evaluation—The linearity evaluation includes a check on the imprecision of the best-fit line or curve. The purpose of this check, formally called a power calculation, is to reduce misclassification errors from poor fit of the regression model. We calculate the specific limit within the statistical algorithm; it depends on the number of included results, the goal for total error, the mean of your included results, and whether your results are fit by a line or curve.

Acceptable Nonlinearity Limits for Linearity Evaluation—In the linearity evaluation, small deviations from linearity are acceptable if, overall, the nonlinearity is within a clinically acceptable range. We calculate the specific limit within the statistical algorithm; it depends on the number of included results, the goal for total error, the mean of your included results, and the shape of the best-fit curve (quadratic or cubic) used to detect nonlinearity. (See definition of best-fit curve.)

Allowable Error for Calibration Verification—The allowable error is the larger of the two limits listed on the evaluation report and in the Participant Summary. Specification of an alternate allowable error for low values prevents the error limits from becoming too small at low concentrations.

Allowable Error for Diluted/Extended Linearity Evaluations—The allowable error for extended range specimens and diluted specimens show the acceptable difference between your results and the value predicted from the best-fit line. If the mean of your results is within the allowable error limit, we will extend the linear range to include consecutive extended range or diluted specimens.

Analytical Measurement Range (AMR)—The analytical measurement range is the range of analyte values that a method can directly measure on the specimen without any dilution, concentration, or other pretreatment not part of the usual assay process.[1]

Best-fit Curve—The first step in the CAP linearity evaluation evaluates a series of polynomial curves fit to your data. This is generally known as the polynomial method.[6, 7] In this step, we may select a quadratic or cubic curve as the best descriptor of the relationship between your reported results and the relative analyte concentration. In a second step, we assess the amount of nonlinearity against a non-zero threshold. You will receive an evaluation of Nonlinear and the best-fit curve is plotted in Linearity Plot 1 if the amount of nonlinearity exceeds this non-zero threshold.

Best-fit Line—The best-fit line is the straight line fit to your included results using linear regression.

Best-fit Targets—The best-fit targets are the points on your best-fit line corresponding to each specimen for which you reported results.

Excluded Results—We exclude assay results from the calculation of your best-fit line for one of two reasons: 1) Your evaluation is not linear in the full range of reported results; or 2) results are evaluated with the diluted/extended range protocol.

Extended Range Specimen—If the relative concentrations have a large gap in the admixture ratios that could permit some specimens to have severely unequal influence on the fitted regression line, we will identify one or more of the specimens as an extended range specimen and use the diluted/extended linearity evaluation.

Goal for Total Error (TE)—Total error is the net or combined effect of random and systematic errors in a test system.[3] We use one-quarter of the goal for total error plus sampling variability estimated as part of the statistical analysis as the limit of acceptable nonlinearity in the linearity evaluation.

Imprecision of Best-fit Line or Curve—Imprecision in the linearity evaluation refers to the differences between your reported results and the best-fit line or curve. One component of this quantity is an estimate of your repeatability or within-run precision, and it also includes a contribution from how well the best-fit line or curve matches your assay means. The imprecision of your best-fit line or curve could be due to unacceptably large differences between one or more pairs of assays or large differences between your assay means and the best-fit line or curve.

Included Results—Included results are assay results that are included in the calculation of the best-fit line or curve in the linearity evaluation.

Insufficient Data for Evaluation—Depending on the Survey, you are required to have four or five consecutive pairs of results to receive a linearity evaluation. If you have at least three undiluted consecutive pairs of results, you will receive a calibration verification evaluation even if you have insufficient data for the linearity evaluation. Your Executive Summary will indicate if you have not reported data for the minimum required number of specimens.

Peer Mean—The peer mean is the mean calculated from peer values at each specimen after outliers are removed from the data.

Relative Concentration—The relative concentration values are the rescaled values for a range of concentration levels where the lowest concentration is set to zero and the highest is set to one. We have selected this scale to reflect the common use of admixtures to produce specimens. High and low specimen pools are mixed in varying proportions to produce a set of specimens with known relationships among the concentration or activity levels. A value of zero means the mixing proportion is zero from the high specimen pool and, by subtraction, one from the low specimen pool. A value of 0.25 means the mixing proportion is 0.25 from the high pool and 0.75 from the low pool. A value of one means a mixing proportion of one from the high specimen pool and zero from the low specimen pool.

Repeatability—Repeatability is the measurement precision of replicate measurements on the same or similar objects over a short period of time. Repeatability may also be called within-run precision in laboratory medicine.

Reportable Range—The reportable range is the span of test result values over which the laboratory can establish or verify the accuracy of the instrument or test system measurement response. (Definition provided in Current CLIA Regulations, Subpart A, Section 493.2)

Frequently Asked Questions

Interpretation of the CVL Evaluations

Q: On Linearity Plot 1, it looks like my results are very close to the line. Why is my evaluation result Imprecise (Poor Repeatability and/or Fit) or Nonlinear?

A: Linearity Plot 2 shows the differences between your results and the best-fit line on a scale that makes it easier to see differences that are large relative to the error goals for a particular analyte. If your evaluation is Imprecise (Poor Repeatability and/or Fit) or Nonlinear, it indicates that the overall measure of differences is larger than the allowable error limit. Large deviations between assay replicates (poor repeatability) and/or large deviations between replicate means and the best-fit line (poor fit) contribute to results of Imprecise (Poor Repeatability and/or Fit) or Nonlinear.

Q: All of my mean values are very close to the peer mean values. How can my linearity evaluation be Nonlinear?

A: With the exception of the Blood Gas Survey (LN13), we assess linearity using the relative concentration or activity, not the peer mean values. The mixing ratios used during the specimen manufacturing process determine the relative concentrations. If your peer group means were nonlinear in relation to the relative concentrations and your linearity results were similar to others in your peer group, then your results would also be Nonlinear.

Q: How are the calibration verification allowable errors and goals for total error determined?

A: The Instrumentation Resource Committee (IRC) determines the total error goals for each new analyte, and they are periodically reviewed by this committee. In this process, the IRC reviews both published error goals based on biological variation and the peer group performance statistics from relevant CAP proficiency testing and CVL Surveys.

Q: Why do the calibration verification and linearity evaluations use different error limits?

A: The calibration verification evaluation is intended to have more stringent error limits than proficiency testing to allow early identification of analytical error. Also, because this evaluation uses means of replicate measurements, it is appropriate to consider some reduction in the error limits. The use of **one-half the total error goal** reflects a reduction in error limits due to the use of means.

For the linearity assessment, participants take duplicate measurements of multiple samples within the same run. This precision component is expected to be similar in magnitude to within-run precision. The linearity imprecision limit has been set as a small fraction, **approximately one-quarter,** of the total error goal because many sources of variation are excluded from this assessment. The use of one-quarter the total error goal aligns with well-established recommendations to use performance goals that will prevent failure on proficiency testing challenges.[9]

Q: On my linearity evaluation report, the best-fit target for my first specimen is "<0." How can the best-fit target be a negative number?

A: As the best-fit line is influenced by all data points, this could be due to random measurement error across the range of specimen values. It could also be an indication of nonlinear recovery at the low end.

Q: I received a diluted/extended linearity evaluation, but it does not include the dilution specimen analysis and my Linearity Plot 2 does not show the allowable error limits for my diluted specimen. Why have these elements been omitted from my report?

A: We evaluate undiluted specimen results first, using the standard linearity evaluation. If your undiluted results are not linear in the full range, we do not evaluate your diluted results. Since your diluted results are not evaluated, we do not calculate the allowable error limits for the diluted specimen analysis or include those limits in Linearity Plot 2.

Q: Why is there a limit for the difference between Assay 1 and Assay 2 in the diluted/extended linearity evaluation?

A: Large differences between replicate measurements may be an indication of poor repeatability. Problems with measurement precision need to be addressed in order to obtain a meaningful assessment of linearity.

Q: Why is there a diluted/extended linearity evaluation for extended range specimens?

A: When there are severely nonequidistant relative concentrations, the specimens with higher concentration have large influence on the fitted regression line and allowable error estimates. The extended range linearity evaluation reduces the chance of “false linear” misclassifications from extreme nonequidistant admixture ratios.

Q: Why does my Linearity Plot 2 have points outside the grey shaded box, even though my evaluation result is Linear over the full range of my results?

A: Points can fall outside of the shaded area for two reasons: 1) Since we use an average to estimate imprecision, many small differences can offset a few larger differences, and 2) clinically insignificant nonlinearity (curved fit) can contribute to differences between your results and the best-fit straight line. In the first case, larger differences between replicates may be an early warning sign of poor repeatability over a subset of your linear range. In the second case, evidence of a U-shaped or inverted U-shaped trend in the difference plot may be an early warning sign of nonlinearity.

Method Verification and Regulatory Requirements

Q: How often is AMR verification required?

A: As specified in the CAP LAP Chemistry and Toxicology Checklist, laboratories are required to verify the AMR at least every six months, or if changes are made to an assay without recalibration. If your laboratory performs calibrations or calibration verification at least every six months using material that spans the low, mid, and high range of the AMR, separate AMR verification is not required. Coagulation tests based on direct measurement of an analyte require AMR verification at least every six months according to the LAP Hematology and Coagulation Checklist.

Q: How often is calibration verification required?

A: If your laboratory performs calibrations at least every six months, current CLIA regulations and the CAP Laboratory Accreditation Program do not require separate calibration verification. Otherwise, laboratories are required to confirm calibration verification at least every six months, or if changes are made to an assay without recalibration.

Q: What am I required to do if my linearity evaluation result is Nonlinear and my calibration verification result is Verified?

A: If your evaluation is Nonlinear, but your results are similar to your peer group, the nonlinearity may be inherent in the method and it is not specific to your instrument. The same scenario may occur due to a nonlinear matrix effect in the Survey material. In this case, it may be useful to investigate the apparent nonlinearity with previously tested patient specimens to identify a possible matrix effect. If your CAP calibration verification evaluation results include low, midpoint, and high values and the evaluation criteria are acceptable to your laboratory, your AMR is verified.

Q: What am I required to do if my linearity evaluation result is Linear but my calibration verification result is Different?

A: Make sure you are reporting in the correct units of measure and have not made a decimal place error. If your results are in agreement for some specimens, look for evidence of a peer-level problem. The peer group means may be influenced by laboratories that are nonlinear due to instrument differences or because of aging reagents. In this case, the coefficients of variation (CV) for the higher specimens tend to be much larger than the CVs of the midrange specimens. If your evaluation is Linear and there is evidence that the mean of your peer group includes nonlinear results, you should note that you suspect a peer-level problem.

Q: What is required by the LAP checklists to demonstrate linearity?

A: The LAP Chemistry and Toxicology Checklist requires both calibration verification and AMR verification. You can use successful performance in the CVL Surveys to verify your AMR, but you must confirm that the evaluation includes specimens near the low, midpoint, and high values of the AMR. The LAP Hematology and Coagulation Checklist states that coagulation tests based on direct measurement of an analyte require AMR verification. Calibrated tests that directly measure activity or concentration of an analyte by enzyme immunoassay, immunoturbidity, or chromogenic methods require AMR verification. Clot-based tests do not require AMR verification. The checklist also states that linearity studies are not required for calibration and calibration verification of CBC instruments.

Q: Is it more important to be Linear or Verified?

A: If poorly performing laboratories influence peer means, a linear result is a direct confirmation that your laboratory is performing well across the measuring interval. In this case, you should give your linearity evaluation priority over calibration verification. However, sometimes there are nonlinear matrix effects in the Survey material. In this case, if you suspect a peer-level failure of linearity, you should give more weight to your calibration verification evaluation.

Q: What other options do I have for AMR verification if my analytical measurement range is not covered by the CAP material?

A: If you are using the CVL Survey to fulfill regulatory requirements, you may need to find an alternative product that has higher (or lower) specimens, utilize residual patient specimens for AMR verification, or adjust your AMR to reflect the range that can be successfully verified.

Q: What am I required to do if only three specimens are within my analytical measurement range?

A: You may report three undiluted sets of results and receive a calibration verification evaluation. You may also report additional diluted results and receive both a linearity and calibration verification evaluation. See the kit instructions for additional information on reporting results beyond your AMR.

Data Processing and Reporting

Q: What is an LN *Express* evaluation and how does it benefit me?

A: For the majority of the CVL Surveys, we offer the LN *Express* service, which provides your linearity evaluation via e-*LAB* Solutions within two business days of receipt of your data. Results may be submitted by any method, but they must be received by the due date printed on your result form.

The LN *Express* evaluation provides the opportunity to verify your submitted results, review your linearity evaluation, and submit any necessary data corrections prior to the start of the CVL Survey processing. If data revisions are submitted prior to the survey due date, we will generate another LN *Express* evaluation.

Your complete report package, including the calibration verification evaluation, will be generated after the survey due date when the peer summary results can be calculated.

Q: Why does it take longer to get a calibration verification evaluation and Participant Summary report?

A: We always use data from the current mailing to calculate peer-based target values and peer group summary statistics. We strive to include as much data as possible, especially for smaller peer groups. We believe this provides the most accurate peer group information for evaluations that utilize peer-based target values. Once we have begun summarizing the peer group data, it takes several days to generate and review the evaluation reports. We spend additional time formatting the Participant Summary reports, which may include commentary by Instrumentation Resource Committee members. Please note that you will have faster access to your full evaluation report package with e-*LAB* Solutions.

Appendix 1: Worksheet for Calibration Verification Self-Evaluation

Use this worksheet to complete a calibration verification self-evaluation for any number of specimens. If you have results near the low, middle, and high end of your AMR and the allowable error limits meet your laboratory's criteria, your self-evaluation may be sufficient to fulfill regulatory requirements for calibration verification or AMR verification.

Survey and Analyte: _____

Allowable Error: _____% or _____
(A) (B)

	(1)	(2)	(3)	(4)	(5)
Specimen Number	Your Mean	Target Value	Difference	Percent Difference	Within Limit (Yes/No)

(6) **Result:**_____

(1) Record the mean values of replicate assays for each specimen.
(2) Record the corresponding target value, ie, peer mean, or an accuracy-based assigned value.
(3) Difference = column (1) – column (2). If the absolute value of this result is less than the Allowable Error (B), record a "Yes" in column (5) and skip column (4).
(4) Percent Difference = column (3) / column (2) × 100. If the absolute value of this result is less than the Allowable Error (A), record a "Yes" in column (5).
(5) If both the Difference and Percent Difference are greater than the Allowable Error, record a "No" in column (5). Note that both differences can be positive or negative, so compare the absolute value of the differences with the allowable error.
(6) Calibration is Verified over the full range of specimens if all results in column (5) are "Yes." The calibration evaluation is Different if there are any "No" results in column (5).

Example: LN2 Lipase U/L

Allowable Error: 17.5 % or 8 U/L
(A) (B)

	(1)	(2)	(3)	(4)	(5)
Specimen Number	Your Mean	Target Value	Difference	Percent Difference	Within Limit (Yes/No)
LN-01	20.5	22.7	2.2		Yes
LN-02	103.5	99.2	4.3		Yes
LN-03	182.0	171.1	10.9	6.4%	Yes

(6) **Result:** Verified

Appendix 2: CVL Survey Investigation Checklist for Problematic Results

Survey Name: _____ **Evaluation Date:** _____

Analyte: _____

Specimen Handling	YES	NO	NA
Was testing material received in the laboratory within an appropriate time after shipment?	☐	☐	☐
Were Survey specimens stored as indicated in the kit instructions?	☐	☐	☐
Were any special instructions provided in the kit instructions performed as indicated?	☐	☐	☐
Were the Survey specimens mixed adequately before sampling?	☐	☐	☐
Were the correct tests performed on the correct vial of testing material?	☐	☐	☐

A response of "No" to any of these questions may indicate a specimen handling error. These types of errors could indicate a failure to read the instructions provided with the Surveys materials.

Analytical	YES	NO	NA
Was the written procedure followed?	☐	☐	☐
Was instrument maintenance performed on schedule?	☐	☐	☐
Were quality control results acceptable?	☐	☐	☐
Was the most recent calibration acceptable and within established stability limits at the time testing was performed?	☐	☐	☐
Does a review of recent proficiency testing results or past CVL results indicate evenly distributed data without bias?	☐	☐	☐
Were the reagents prepared according to procedure?	☐	☐	☐
Were the reagents within their open stability acceptable range?	☐	☐	☐
Was the intended result within the measuring range for the instrument?	☐	☐	☐
Was the dilution protocol followed when diluting samples that are out of range?	☐	☐	☐
Does a review of records indicate that there were no related instrument/test problems noted prior to or after the testing was performed?	☐	☐	☐

A response of "No" to any of these questions may indicate an analytical error. These types of errors could indicate a failure to follow recommended instrument maintenance and calibration. You may need to review the instructions provided with the testing material and/or laboratory procedures. If recalibration has not already occurred, recalibrate the instrument.

Clerical	YES	NO	NA
Were the results correctly transcribed from the instrument read-out or report?	☐	☐	☐
Was the correct instrument/method/reagent code reported on the result form?	☐	☐	☐
Do the units of measure match between the result form and the instrument results?	☐	☐	☐
Is the decimal place correct?	☐	☐	☐
Does the submitted result match the result found on the calibration verification evaluation report?	☐	☐	☐
If the result was out of range and a dilution was performed, was the correct dilution factor used in the calculation of the final result?	☐	☐	☐

A response of "No" to any of these questions may indicate a clerical error. Although reporting of testing results is unlike those for patient results, clerical errors may indicate a need for additional staff training, review of kit instructions, or investigation of the reporting format provided by the testing device. If results reported on the result form do not match the results found on the evaluation report, please contact the CAP Customer Contact Center at 800-323-4040.

Corrective Action:

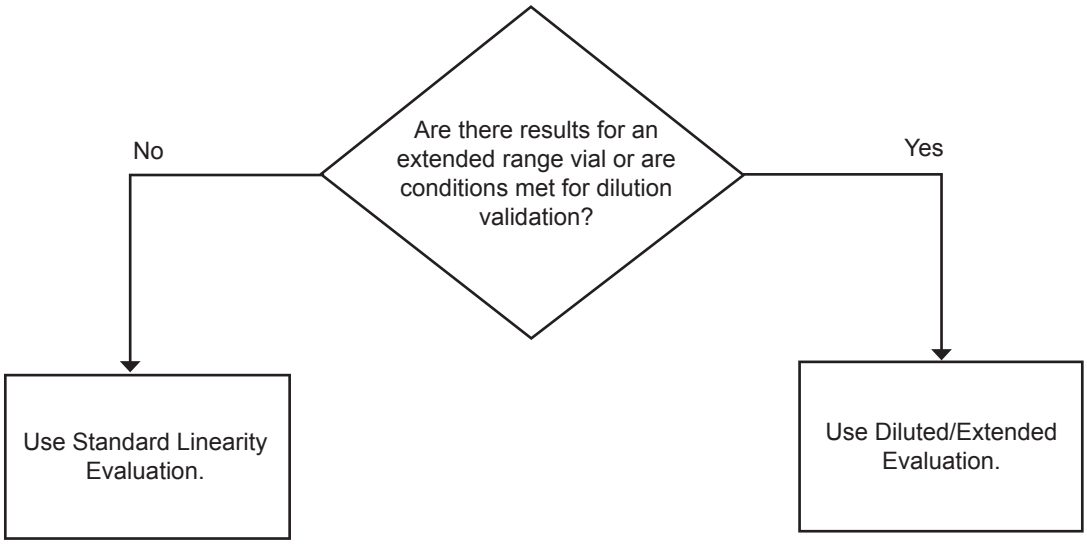
Additional Notes:

Investigated by: _____ **Date:** _____

Reviewed by: _____ **Date:** _____

Appendix 3: Flow Charts

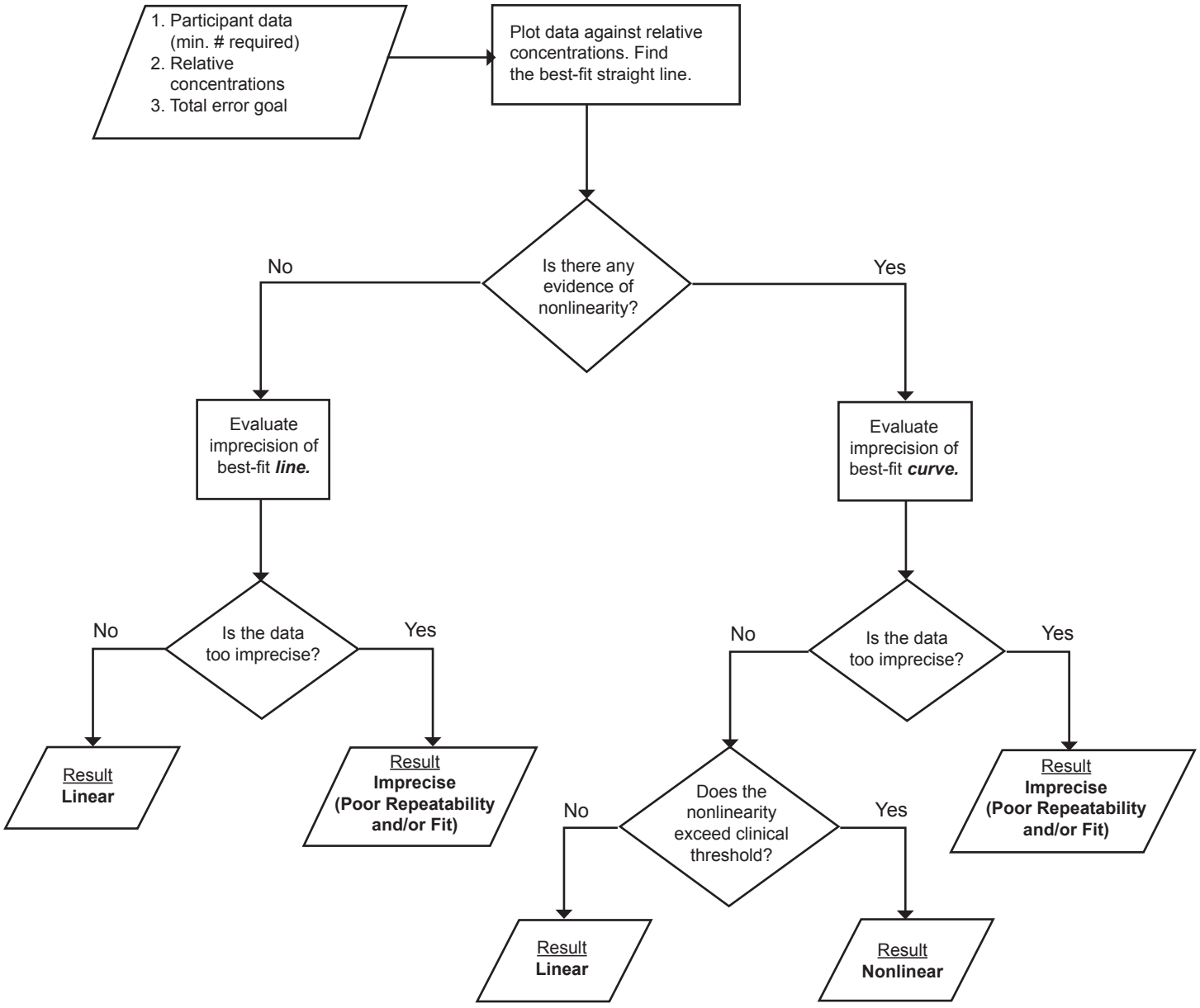
Flow Chart 1: Selection of Linearity Evaluation Type



Conditions for extended range:
1. The CAP has identified the specimen as an extended range specimen.
2. Participant submits results for at least four additional specimens.

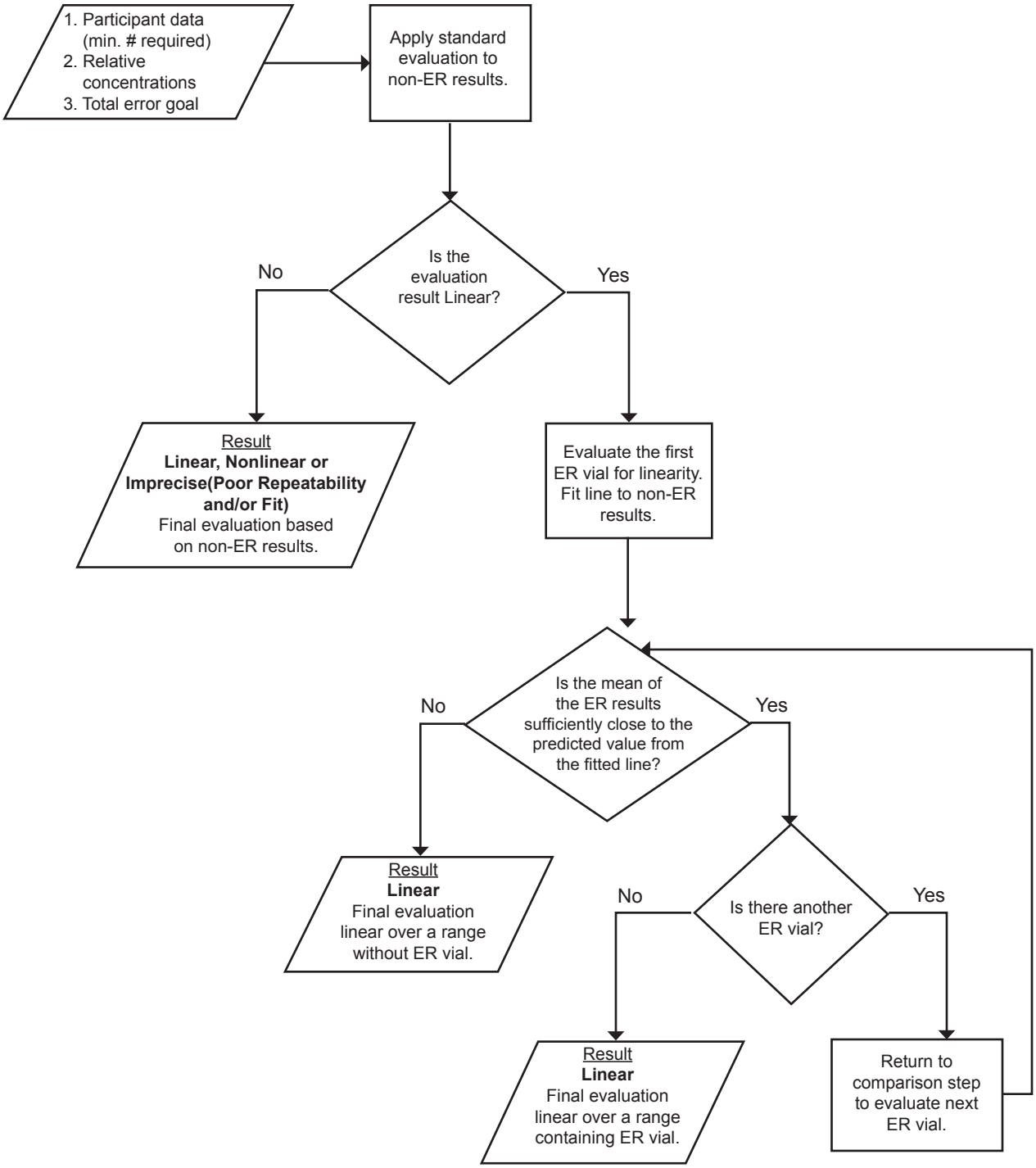
Conditions for dilution validation:
1. At least four consecutive specimens must be undiluted.
2. Both replicates must be diluted.

Flow Chart 2: Standard Linearity Evaluation

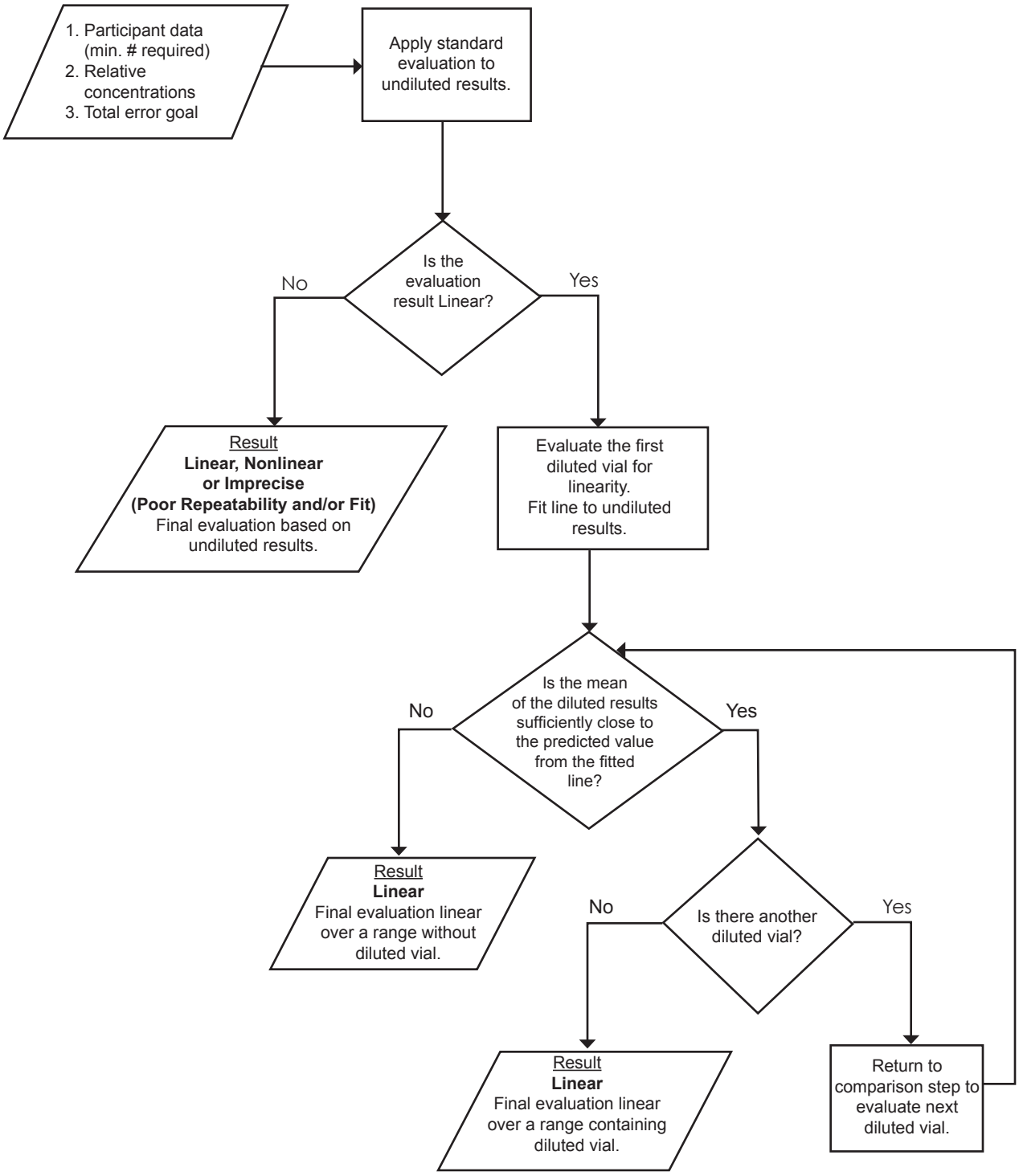


Note: If the result is ***Nonlinear*** *or* ***Imprecise (Poor Repeatability and/or Fit)****, an evaluation may be rerun on a subset of your data.*

Flow Chart 3: Diluted/Extended Linearity Evaluation When Results Include Extended Range (ER) Specimens



Flow Chart 4: Diluted/Extended Linearity Evaluation When Results Include Diluted Specimens



Appendix 4: Technical Guide to the CAP Linearity Evaluation

This appendix is intended for participants who have had training equivalent to a second course in statistical methods and would like more details about the CAP linearity evaluation. Because this information is intended for participants with some training in statistical methods, we do not attempt to provide complete details for carrying out the calculations. You can reference any statistical text for more information about fitting linear and polynomial regression models and evaluating the statistical significance of regression coefficients (steps 1 and 2). The calculations noted in steps 3 and 4 require more advanced techniques. While these are fully documented in the manuscript by Kroll, et al in the September 2000 issue of Archives of Pathology & Laboratory Medicine*, the explanation in this appendix will provide a set of approximate calculations using standard regression output.*

As a preliminary step, determine the type of evaluation that has been or will be used—Standard or Diluted/Extended. See the linearity flow charts in Appendix 3 for reference. If you are trying to duplicate results on your linearity evaluation report, you must match the range of included specimens. In a Diluted/Extended evaluation, you will use the standard evaluation for your undiluted / nonextended range specimens (steps 1 through 4 below).

1. Plot the measured results against the relative concentration and find the best-fit straight line using linear regression.

Plot your results on the y-axis and the relative concentrations on the x-axis. The linearity evaluation report lists the relative concentrations for each analyte. You can use any statistical software to fit the regression line.

2. Evaluate the data for evidence of nonlinearity.

Nonlinearity is estimated using polynomial regression. If there are two results reported for at least five test specimens, fit a cubic polynomial and assess the cubic term using a standard test of statistical significance. If there are results for only four test specimens, or if the cubic term is not statistically significant, fit a quadratic polynomial. Assess the quadratic term using a standard test of statistical significance. You can use any standard regression software to fit the polynomials and test the significance of the cubic and quadratic terms. The CAP uses a significance level of alpha = 0.05 for this evaluation. If there is no evidence of nonlinearity using the polynomial method, the best-fit curve is the straight line that was found in step 1. Otherwise, the best-fit curve is either the quadratic or cubic polynomial. To evaluate the imprecision in step 3, you need to identify the appropriate best-fit line or curve.

Note that in step 2, the hypothesis tests use a threshold of zero to detect deviations from linearity (ie, the Null Hypothesis in each statistical test is that the nonlinear coefficient equals 0). After screening for imprecision of the best-fit line or curve, the CAP linearity evaluation assesses the amount of nonlinearity described by the quadratic or cubic polynomial relative to a non-zero threshold, described in step 4. For step 4, you will need to calculate the fitted values from both the best-fit straight line and the best-fit curve from your regression results.

3. Evaluate the imprecision of the best-fit line or curve.

We evaluate imprecision based on the differences between the participant results and the best-fit curve or line. If the differences exceed the limit on imprecision, we conclude that the data are inadequate to enable a meaningful linearity evaluation, either because of poor repeatability or poor conformance to the specified linear and polynomial models. The limit on the imprecision around the best-fit line or curve is determined by the goal for total error for the analyte, the number of participant results, and a calculated value derived from the formal statistical evaluation. While the exact calculation depends on several input variables, the limit on the imprecision is always slightly larger than one-fourth of the goal for total error. If you are performing a self-evaluation, using a multiple of 0.25 times the goal for total error provides an absolute lower bound for the imprecision screen. Alternatively, you can use a multiplier from the table below determined by the number of included results, for a close approximation to the CAP evaluation.

To calculate the estimated imprecision: Find the estimated standard error of the best-fit line or curve from steps 1 and 2. The software that you use to fit your regression models will calculate the standard error of the best-fit line or curve.

To calculate the imprecision limit: You need to know the mean of your results included in the regression model, the number of results included in the model, the analyte's goal for total error, and the imprecision multiplier from Table 1. The limit on imprecision is the square root [(number of included results)/(multiplier from Table 1)] x [goal for total error x .25] x (mean of included results).

Table 1—Imprecision multiplier

Best fit from step 2	Multiplier
Line	6.3
Curve – quadratic polynomial	6.3
Curve – cubic polynomial	6.5

Your evaluation result is Imprecise (Poor Repeatability and/or Fit) if your best-fit line or curve fails the imprecision screen. If you have no evidence of nonlinearity from step 2 and the results for your best-fit line pass the imprecision screen, your evaluation result is Linear. If you found evidence of nonlinearity in step 2 and your results for your best-fit curve pass the imprecision screen, proceed to step 4.

4. Evaluate nonlinearity relative to a non-zero clinical threshold, if applicable.

We use the difference between the best-fit curve from step 2 and the best-fit straight line from step 1 to evaluate the clinical relevance of nonlinearity. If the summary measure of the difference exceeds the clinical threshold, the evaluation result is Nonlinear. Otherwise, the evaluation is Linear.

The summary measure of the difference between the best-fit curve and best-fit line is called the average deviation from linearity (ADL). We use a multiple of the goal for total error to determine the threshold of clinical importance. Note that this evaluation of nonlinearity permits small, clinically unimportant, deviations from linearity. The evaluation in step 2 flagged any apparent deviation from nonlinearity.

To calculate the ADL: Calculate a single set of linear fitted values from the best-fit line from step 1. Note that if you generate fitted values in your regression software, you may have two sets of identical fitted values. Next, calculate the nonlinear fitted values from the quadratic or cubic curve that was fit in step 2. Again, use just one fitted value for each specimen level. Calculate the squared differences = (linear fits – nonlinear fits).[2] Calculate the average squared difference = (sum of the squared differences) / (number of specimen levels). Finally, the ADL = the square root of the average squared difference.

To calculate the ADL limit: The ADL limit is the 95th percentile of a noncentral Chi-square distribution. Because we feel this calculation is beyond the level of even a second course in statistics, we refer readers with advanced training to the September 2000 issue of *Archives of Pathology & Laboratory Medicine* for a complete description of how to calculate the exact ADL limit. The ADL limit is always slightly larger than one-fourth the goal for total error, so we recommend using this as an approximate limit. Specifically, to evaluate the significance of the nonlinearity estimated in the polynomial method using a non-zero threshold, compare the calculated ADL to 0.25 x (goal for total error) x (mean of included results).

5. Evaluate diluted or extended range specimens, if applicable.

In the Diluted/Extended evaluation, we evaluate your diluted or extended range specimens only if the full range of undiluted or non-ER results are linear. When these conditions are met, you can extend the best-fit straight line from step 1 through the diluted or extended range vials using the relative concentrations listed on your linearity evaluation report. The difference between the observed and expected results is equal to the mean of your replicates minus the extrapolated value from the best-fit line.

Your linearity evaluation page shows the allowable limits for the differences between the observed and fitted results used in the CAP evaluation. The allowable error limit equals the best-fit target times one-half the goal for total error. The evaluation also includes a screen for large replicate differences. If the difference between replicates for the diluted or extended range specimen exceeds the best-fit target times one-third the goal for total error, the linear range will not be extended, regardless of the agreement between your mean and the best-fit target.

If you would like to override the CAP evaluation, you may determine your own limits for the allowable differences. If you have specific questions about the calculation of the limits in the CAP evaluation, we recommend that you contact the CAP Customer Contact Center for more information.

References

1. College of American Pathologists, Commission on Laboratory Accreditation. Chemistry and Toxicology Checklist. Northfield, IL: CAP, 2012.
2. Current CLIA Regulations. Centers for Disease Control and Prevention website. cdc.gov/clia/regs/toc.aspx. Updated January 24, 2004. Accessed September 29, 2012.
3. Linnett K, Boyd JC. Selection and analytical evaluation of methods—with statistical techniques. In Burtis CA, Ashwood ER, Bruns DE, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. St. Louis, MO: Elsevier Saunders; 2006.
4. Clinical and Laboratory Standards Institute/NCCLS. Evaluation of the linearity of quantitative measurement procedures: A statistical approach; Approved guideline. CLSI/NCCLS Document EP6-A. Wayne, PA: Clinical and Laboratory Standards Institute; 2003.
5. Kroll MH, Praestgaard J, Michaliszyn E, Styer PE. Evaluation of the extent of nonlinearity in reportable range studies. *Arch Pathol Lab Med*. 2000;124(9):1331–1338. http://arpa.allenpress.com.
6. Kroll MH, Emancipator K. A theoretical evaluation of linearity. *Clin Chem*. 1993;39(3):405–413.
7. Emancipator K, Kroll MH. A quantitative measure of nonlinearity. *Clin Chem*. 1993;39(5):766–772.
8. College of American Pathologists, Commission on Laboratory Accreditation. Hematology and Coagulation Checklist. Northfield, IL: CAP, 2012.
9. Burnett RW, Westgard JO. Selection of measurement and control procedures to satisfy the Health Care Financing Administration requirements and provide cost-effective operation. *Arch Pathol Lab Med.* 1992;116(7):777–780.

Notes

COLLEGE of AMERICAN PATHOLOGISTS

325 Waukegan Rd.
Northfield, IL 60093
800-323-4040 | cap.org