SPEX CertiPrep App Note

In Pursuit of the True Value: Error and the Use of Standards in Producing Accurate Data

Introduction

Analytical laboratories face more challenges and regulations than ever before as accreditation bodies issue increasing numbers of guidelines; and regulatory agencies increase the number of analytes that need to be reported while the levels of detection required decrease. A lot of time, effort and money is invested in deciphering the data and determining its validity and accuracy. Often terms which describe data are used incorrectly or interchangeably to try to validate a data set or methodology (i.e. error vs. uncertainty, precision vs. accuracy, etc). One of the first steps to understanding and validating data is the proper application of appropriate statistics and the understanding of the use and terminology of analytical processes.

True Value, Accuracy and Precision

All analytical laboratories pursue 'good' data and 'true' values. The reality is that true values are never absolute. The nature of a true value is that it, in itself, contains uncertainty and error which make them somewhat indeterminate. True values are obtained by perfect and error-free measurements which do not exist in reality. Instead, the expected, specified or theoretical value becomes the accepted true value. Analysts then compare the observed or measured values against that accepted true value to determine accuracy or 'trueness' of the data set.

Often accuracy and precision are used in the same context when discussing data quality. In reality, accuracy and precision are very different assessments of data and the acquisition process. Accuracy is the measurement of individual or groups of data points in relationship to the 'true' value. In essence, accuracy is how close your data gets to the target and is often expressed as either a form of numerical or percent difference of the observed result and the target or 'true' value.

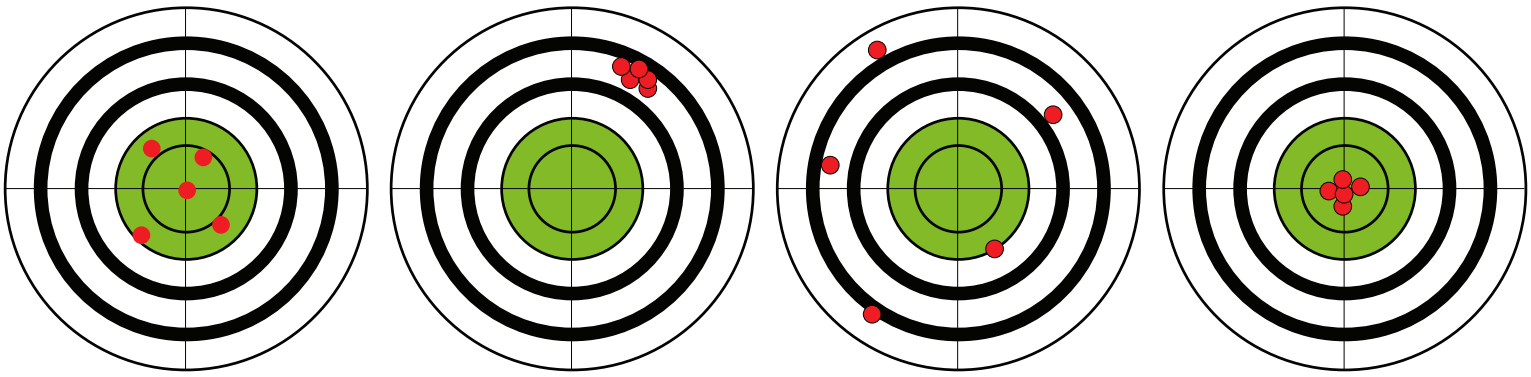

Figure 1. Representations of Accuracy and Precision

Precision, on the other hand, is the measurement of a data set for how well the data points relate to each other. It is the measure of how clustered the data points fall within the target range and is often expressed as a calculation of standard deviation of the data in some form. Precision is an important tool for the evaluation of instrumentation and methodologies by determining how data is produced after varied replications.

Repeatability and reproducibility measure the quality of the data, method, or instrumentation by examining the precision under the same (minimal difference) or different (maximal difference) test conditions. Repeatability (or test-retest reliability) is the measurement of variation arising when all of the measurement conditions are kept constant such as the same location, the same procedure, the same operator, the same instrument run under the same conditions run in repetition over a short period of time. Several standards organizations, such as ASTM, define the parameters of repeatability, intermediate precision and reproducibility in order to create and publish test methods (see Table 1).

Table 1. Conditions for Precision (ASTM E177 & E456)

	Repeatability	Intermediate Precision	Reproducibility
Laboratory	Same	Same	Different
Operator	Same	Different	Different
Apparatus	Same	Same in type or actual apparatus	Different
Time Between Replicates	Less than a day	Multiple days	Not specified

Reproducibility is the measurement of variation arising in the same measurement process occurring across different conditions such as location, operator, instruments, and over long periods of time.

Another way of looking at accuracy and precision is in terms of measurement of different types of error.

Estimating Error

The most common misconception regarding analytical data and results revolves around the concepts of error, mistakes and uncertainty. In general, an error is a deviation or difference between the estimated or measured value and the true, specified, or theoretically correct value. If accuracy is the measurement of the difference between a result and a 'true' value, then error is the actual difference or the cause of the difference. The estimation of error can be calculated in two ways, either as an absolute or relative error. Absolute errors are expressed in the same units as the data set and relative errors are expressed as ratios such as percent, fractions, etc.

Absolute accuracy error is the true value subtracted from an observed value and is expressed in the same units as the data. For example, if a stated expected true value of an analysis is 5 ppm but the resulting value is 6 ppm, then the absolute error for that data point is 1 ppm. Relative accuracy error is the true value subtracted from the observed value and the result is divided by the true value. Errors in precision data are most commonly calculated as some variation of the standard deviation of the data set. An absolute precision error calculation is based on either the standard deviation of a data set or values taken from a plotted curve. A relative precision error is most commonly expressed as relative standard deviation (RSD) or coefficient of variance (CV or %RSD) of the data set (see Table 2).

Table 2. Accuracy & Precision for Absolute and Relative Errors

	Absolute	Relative
Accuracy	$E_{abs} = X_o - X_t$	$E_{rel} = X_o - X_t / X_t$
Precision	δ of data set or value taken from a curve	RSD or CV of data set

Types of Error

There are many types of error associated with scientific and statistical analyses. The most common errors, in regards to data, are observational or measurement errors which are the difference between a measured value and its true value. Most measured values contain an inherent aspect of variability as part of the measurement process which can be classified as either random or systematic errors.

Random (or indeterminate) errors lead to measured values which are inconsistent with repeated measurements. Systematic (or determinate) errors are introduced by inaccuracy from the measurement process or analytical system. There are some basic sources for systematic error in data. These sources are: operator or analyst, apparatus and environment, method, or procedure. Systematic errors can often be reduced or eliminated by observation, record-keeping, training, and maintenance. Operator or analyst errors can occur due to inattentiveness, lack of training or misinformation. Operator or analyst errors are most often called mistakes. Apparatus or laboratory environment errors can occur with improper maintenance, substandard laboratory environment and materials (i.e. improper volumetrics, improper calibration, poor environmental temperature and humidity controls, etc.). Method or procedure errors can occur with poor method validation or lack of periodic updates as equipment or materials change.

In cases where systematic errors lead to results in a data set that trend higher or lower than the 'true' value, the difference is considered to be a bias. A positive bias creates a trend where results are higher than the expected value, while a negative bias displays results lower than the expected value. Determinate errors and bias can often either be eliminated by examining the sources for error and correcting the problem or, in some cases of consistent error, may be corrected or adjusted for in the instrumentation or procedure.

Random or indeterminate errors arise from random fluctuations and variances in the measured quantities and occur even in tightly controlled analysis systems or conditions. It is not possible to eliminate all sources of random error from a method or system. Random errors can, however, be minimized by experimental or method design. For instance, while it is impossible to keep an absolute temperature in a laboratory at all times, it is possible to limit the range of temperature changes. In instrumentation, small changes to the electrical systems from fluctuations in current, voltage and resistance cause small continuing variations which can be seen as instrumental noise. The measurement of these random errors is often determined by the examination of the precision of the generated data set. Precision is a measure of statistical variability in the description of random errors. Precision analyzes the data set for the relationship and distance between each of the data points independent of the 'true' or estimated value of the data to identify and quantify the variability of the data.

Accuracy is the description of systematic errors and is a measure of statistical bias which causes a difference between a result and the 'true' value (trueness). A second definition, recognized by ISO, defines accuracy as a combination of random and systematic error which then requires high accuracy to also have high precision and high 'trueness'. An ideal measurement method, procedure, experiment, or instrument is both accurate and precise with measurements which are all close to and clustered around the target or 'true' value. The accuracy and precision of a measurement value is a process validated by the repeated measurements of a traceable reference standard or reference material.

Using Standards and Reference Materials in True Value

A standard is a known or characterized material used to confirm identity, concentration, purity, or quality. Standards are considered either to be primary standards or secondary standards. A primary standard is sufficiently accurate so that it is not compared to or calibrated by other standards. Primary standards are produced by metrological agencies such as NIST (National Institute of Standards and Technology) in the United States. Primary standards are used to calibrate secondary standards which are produced by secondary standards manufacturers.

Reference materials are standards created either by a primary or secondary producer which are used to provide data regarding accuracy and reliability of analytical results. Certified standards, or certified reference materials (CRMs), are materials which have one or more certified values with uncertainty which have been established using validated methods and are accompanied by a certificate.

The certified value is the accepted established value which can be reasonably attributed to the measured value within the range of the stated uncertainty. The uncertainty characterizes the range of the dispersion of values that occurs through the determinate variation of all of the components which are part of the process for creating the standard. Each of the components in the creation of the standard have a calculated uncertainty which then are all combined to create a combined uncertainty associated with the certified value. For example, in the creation of a chemical standard there could be separate uncertainties for all of the volumetric glassware used in the production of the standard as well as uncertainty from the purity of starting material, variations in the balance, temperature of the laboratory, and purity of the solvents. Each uncertainty for individual components is calculated and combines to form the combined uncertainty for the standard. To be clear, the uncertainty listed on a standard certificate is the measured uncertainty for that standard's certified value and not the expected range of results for an instrument or test method. Each test method or instrument carries its own set of uncertainty calculations which determine the accuracy and precision of that analytical method which is independent of the value on the certificate of the standard.

In addition to certified values and uncertainty, a certificate for a certified reference material can contain statements of traceability and stability, meaning that the certified values can be traced to a primary source and the standards are not reactive during normal use. A stabile standard will retain its properties in the expected timescale when maintained in the environmental conditions and used for the purpose intended and outlined on the certificate.

CRMs have a number of uses including: validation of methods, standardization or calibration of instrument or materials, and for use in quality control and assurance procedures. A calibration procedure establishes the relationship between a concentration of an analyte and the instrumental or procedural response to that analyte. A calibration curve is the plotting of multiple points within a dynamic range to establish the analyte response within a system during the collection of data points. One element of the correct interpretation of data from instrumental systems is the effect of a sample matrix upon an instrumental analytical response. The matrix effect can be responsible for either analyte suppression or enhancement. In analysis where matrix can influence the response of an analyte, it is common to match the matrix of analytical standards or reference materials to the matrix of the target sample to compensate for matrix effects.

Different approaches to using calibration standards may need to be employed to compensate for the possible variability within a procedure or analytical system. Internal standards are reference standards that are either similar in character or analogs of the target analytes that have a similar analytical response are added to the sample prior to analysis. In some cases, deuterated forms of the target analytes are used as internal standards. This type of standard allows the variation of instrument response to be compensated for by the use of a relative response ratio established between the internal standard and the target analyte. A second type of internal standard is a standard addition or a spiking standard. In some analyses, the matrix response, instrument response and the analyte response are indistinguishable from each other as the analyte concentration nears the lower limit of detection or quantitation. A target standard can then be added in known concentration to compensate for the matrix or instrument effects to bring the signal of the target analyte into a quantitative range.

External standards are multiple calibration points (customarily four or more points) that contain standards or known concentrations of the target analytes and matrix components. Depending on the type of analytical techniques, linear calibration curves can be generated between response and concentration which can be calculated for the degree of linearity or the correlation coefficient (r). An r value approaching 1 reflects a higher degree of linearity, most analysts accept values of > 0.999 or better as acceptable correlation.

Calibration curves are often affected by the limitations of the instrumentation. Data can become biased by calibration points biased by instruments limits of detection, quantitation and linearity. Limit of detection (LOD) is the lower limit of a method or system at which the target can be detected as different from a blank with a high confidence level (usually over three standard deviations from the blank response). The limit of a quantitation (LOQ) is the lower limit of a method or system which the target can be reasonably calculated where two distinct values between the target and blank can be observed (usually over ten standard deviations from the blank response) (see Figure 2).

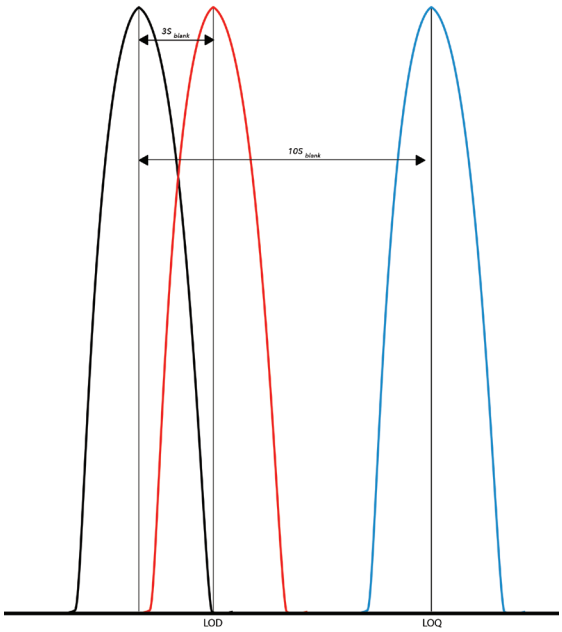

Figure 2. Limits of Detection and Quantitation

A second method of determining levels of detection and quantitation can be considered using a signal-to-noise. Signal-to-noise is the response of an analyte measure on an instrument as a ratio of that response to the baseline variation (noise) of the system. Limits of detection are often recognized as target responses which have three times the response of baseline noise or s/n >/= 3. Limits of quantitation are recognized as target responses which have ten times the response of baseline noise or s/n >/= 10.

Limits of linearity (LOL) are the upper limits of a system or calibration curve where the linearity of the calibration curve starts to be skewed creating a loss of linearity (see Figure 3). This loss of linearity can be a sign that the instrumental detection source is approaching saturation. The array of data values between the LOQ and the LOL is considered to be the dynamic range of the system where the greatest potential for accurate measurements will occur.

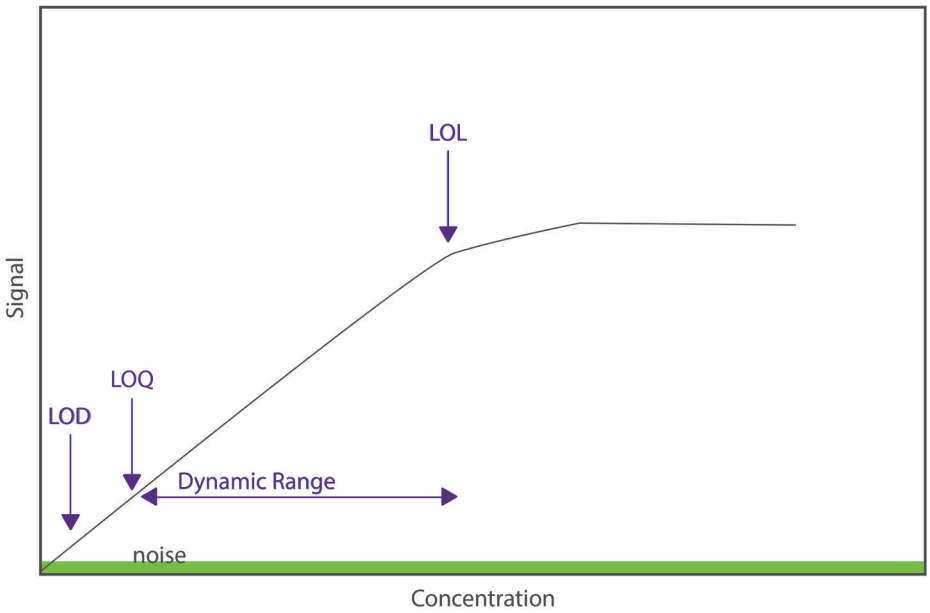

Figure 3. Calibration Curve Limits and Range

The understanding of a system's dynamic range, the accurate bracketing of calibration curves within the range and around the target analyte concentration increase the accuracy of the measurements. If a calibration curve is created that does not potentially bracket all of the possible target data points, then the calibration curve can be biased to artificially increase or decrease the results and create error.

The elimination of error from analytical methods is an ongoing process which forces the analytical laboratory to examine all of their processes to eliminate sources of systematic error and mistakes. It is then a process of identifying the sources of random error in the analysis and sample preparation procedures to calculate the uncertainty associated with each source. Standards and certified reference materials give an analyst the known quantity, character or identity to reference against their samples and instruments to further eliminate error and increase accuracy and precision to bring them closer to their goal of determining the true value.

Phone: 800.LAB.SPEX • 732.549.7144 • Fax: 732.603.9647
CRMSales@spex.com • spexcertiprep.com