INTRODUCTION

A2LA has compiled information for classifying some common types of test methods to meet the A2LA Policy on Measurement Uncertainty for Medical Testing Laboratories. This A2LA Policy is intended to facilitate compliance with ISO 15189 and is subject to change as additional guidance is made available internationally.

This policy was developed by the A2LA Medical Technical Advisory Committee. It provides guidelines for categorizing methods when determining measurement uncertainty. Laboratories must comply with 5.5.1.4 of ISO 15189 regardless of whether a method is listed as Category I, II, or III. This requirement is included below. Table 1 at the end of this document provides guidelines where test methods are grouped by discipline along with their MU category designation.

While the Policy implements only the requirements of ISO 15189, it references concepts in the current edition of ISO/IEC 17025:2017, and the documents referenced therein – namely, ISO/IEC Guide 98-3:*Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)*, ISO 21748:2017 *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty evaluation*, and the ISO 5725 series of documents on *Accuracy (trueness and precision) of measurement methods and results.* The policy is also intended to be consistent with ISO 20914:2019 *Practical Guidance for the Estimation of Measurement Uncertainty.*

Note: P905 – A2LA Metrological Traceability Policy for ISO 15189 Laboratory Testing applies for calibration of reference standards or working standards that require metrological traceability. Calibrations must include a measurement uncertainty that has been calculated in accordance with the Guide to the Expression of Uncertainty in Measurement (GUM).

PRINCIPAL REFERENCE

ISO 15189, clause **5.5.1.4 Measurement uncertainty of measured quantity values**

The laboratory shall determine measurement uncertainty for each measurement procedure in the examination phase used to report measured quantity values on patients' samples. The laboratory shall define the performance requirements for the measurement uncertainty of each measurement procedure and regularly review estimates of measurement uncertainty.

NOTE 1 The relevant uncertainty components are those associated with the actual measurement process, commencing with the presentation of the sample to the measurement procedure and ending with the output of the measured value.
NOTE 2 Measurement uncertainties may be calculated using quantity values obtained by the measurement of quality control materials under intermediate precision conditions that include as many routine changes as reasonably possible in the standard operation of a measurement procedure, e.g. changes of reagent and calibrator batches, different operators, scheduled instrument maintenance.
NOTE 3 Examples of the practical utility of measurement uncertainty estimates might include confirmation that patients' values meet quality goals set by the laboratory and meaningful comparison of a patient value with a previous value of the same type or with a clinical decision value.

The laboratory shall consider measurement uncertainty when interpreting measured quantity values. Upon request, the laboratory shall make its estimates of measurement uncertainty available to laboratory users.

Where examinations include a measurement step but do not report a measured quantity value, the laboratory should calculate the uncertainty of the measurement step where it has utility in assessing the reliability of the examination procedure or has influence on the reported result.

RELATED CONCEPTS FOR DETERMINING THE UNCERTAINTY OF MEASUREMENT

Key references on uncertainty use different terms to describe the tasks associated with determining the uncertainty or quantitative measurements. One important distinction is between the *evaluation* of measurement uncertainty and the *estimation* of measurement uncertainty, both of which are used to *determine* measurement uncertainty. Some key terms are described below as they are used in this Policy.

Evaluation of measurement uncertainty: a process of considering the relevant sources of uncertainty of measurement, the available information, and the user's needs for certainty for their interpretation of the result (adapted from ISO 21748)
Estimation of measurement uncertainty: the process to produce a number for the uncertainty of a particular measurement result; perhaps expressed generally as a percentage of the result (adapted from ISO 21748)
Determination of measurement uncertainty: a combination of evaluation and estimation (adapted from ISO 15189).
Target uncertainty: an objective goal for uncertainty that would make a result sufficiently accurate for its intended use (adapted from ISO Guide 99 on the International Vocabulary for Measurements).

For purposes of clarity, the terms below are used as they are defined in ISO 5725 (which are often misunderstood):
Accuracy
Precision
Bias

LABORATORY PROCEDURE

In this Policy, the laboratory is required to identify and document the applicable measurement uncertainty category (I-III below) for each of the test methods identified on the laboratory's proposed scope of accreditation. Quantitative estimates of measurement uncertainty are not required for Category I methods and uncertainty can be estimated from available data for Category II and Category III methods.

This requirement is in addition to the requirements in ISO 15189 that the laboratory do the following for every measurement method in the Scope of accreditation (summarized from clause 5.5.1.4):

- Determine the measurement uncertainty;
- Define the performance requirements regarding measurement uncertainty;
- Regularly review the estimates;
- Consider measurement uncertainty when interpreting measurement results.
- Estimate components of measurement uncertainty when qualitative results involve a measurement

I. **Test methods that are reported on a qualitative basis, or on a categorical or nominal scale**. These are methods where test items (samples) are classified using visual or microscopic observation or other similar methods to determine, detect, or identify the target. The requirement to calculate measurement uncertainty does not apply to test methods or studies where the end point is an opinion or a diagnosis.
II. **Well-recognized test methods** are those methods that specify limits to the values of the major sources of uncertainty of measurement (usually as considerations in the instructions) and specify the form of presentation of calculated results. This category includes methods or devices approved by the US FDA (this does not include methods that are modified, see Category III, below). There are two significant subclasses of precision statements for these methods:

a) Methods that include statements of precision that are determined with an interlaboratory comparison study and include an estimate of precision of reproducibility (for example, ISO 5725-2).
b) Methods that include statements of precision derived from a single laboratory (perhaps using different instruments), often derived for different concentration (levels) of an analyte.

III **All other test methods**, these include test methods based on published regulatory or consensus methods (examples: CLSI, ISO) that have not been approved by US FDA, and laboratory-developed (or modified) methods, and those test methods needing major (or all) components of uncertainty identified. In such cases measurement uncertainty estimates are to be generated based on appropriate techniques specified below. Laboratory–developed methods require validation per ISO/IEC 17025:2017, section 7.2.2, ISO 5725, or relevant standards for method development (e.g., CLSI, FDA). As part of this validation, the significance of measurement components or the significance of the modifications of the measurement components from the standard test method must be considered so that the measurement uncertainty for the method can be determined.

DETERMINING MEASUREMENT UNCERTAINTY

The laboratory should first evaluate measurement uncertainty by identifying the sources of uncertainty associated with testing technologies and/or test methods. The laboratory should also identify the major components contributing to the uncertainty and, where applicable, present the calculations used for quantifying the measurement uncertainty for the test method. The components of uncertainty should be identified for all test methods or studies, accompanied by reasonable estimates of their magnitude. Then the estimate of the measurement uncertainty may be determined from either reference or control samples, from method validation data, or from combining the individual components.

Category I Test Methods: No calculated estimates of uncertainty are required for test methods that are qualitative, categorical or based on a nominal scale test methods.
Category II Test Methods: For methods in class II a) and b), estimates of uncertainty can be derived from traditional quality control (QC) studies in the user's laboratory, if the QC studies include all major steps in the measurement process (see ISO 20914). This approach requires consideration of uncertainty that can vary by concentration level. If the QC procedure does not include all major steps in the measurement process QC data can still be used if evaluated and revised appropriately, described below. For methods in class II a), estimates of measurement uncertainty can be based on reported estimates of reproducibility if the laboratory can demonstrate that they are competent to use the method (see ISO 21748).

Qualitative and semi-quantitative tests that are based on continuous or quantitative responses and have pre-determined cutoff points are influenced by measurement uncertainty. The effect of the uncertainty can be an incorrect qualitative response. To account for this, many methods have an allowance for an "indeterminate" response. Therefore, samples where results are close to the decision point (if available) are those most at risk, and should be the basis for investigative studies on measurement uncertainty (using, for example, conventional models for detection limits). In these situations, measurement uncertainty can be expressed as either:

1. A traditional MU statement for samples at levels near the decision point(s).
2. A statement about false classification rates for results near the decision point(s).
3. Overall rates of correctness for different known classes of samples (e.g., true positives and true negatives; sensitivity and specificity; etc).

Category III Test Methods: For these methods, MU shall be estimated using available data, published information, and/or designed experiments, as described in the following documents:

- ISO 20914:2019 *Practical Guidance for the Estimation of Measurement Uncertainty*
- A2LA G104: *Guide for the Estimation of Measurement Uncertainty in Testing*

- ISO Guide 98-3:*Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)*
- CLSI EP29: *Expression of measurement uncertainty in laboratory medicine*

DATA SOURCES

Laboratory control sample (LCS) results may be used to estimate MU, provided the samples are an appropriate matrix and concentration. Laboratories should follow the procedures in ISO 20914:2019 *Practical Guidance for the Estimation of Measurement Uncertainty* and/or in ASTM E 2554-18: *Standard Practice for Estimating and Monitoring the Uncertainty of Test Results of a Test Method in a Single Laboratory Using Control Chart Techniques*

Alternatively, they may estimate uncertainty using the following guidance:

1. When the LCS has been through all method steps, then the laboratory can use the standard deviation (SD_P) from the LCS intermediate precision data as an estimate of combined standard uncertainty. A relative SD (or CV) may also be used.
2. When the LCS have not been run through all method steps, then the laboratory should incorporate any appropriate additional components or considerations in the uncertainty calculations, for example, those uncertainty components from sub-sampling, aliquoting or sample preparation. The additional components should be combined with $\mathbf{SD_P}$ using the root sum square **(RSS)** method.
3. When a method has a known consistent bias that is inherent to the method (e.g. low recovery on difficult analytes) the bias must not be added to the uncertainty calculations. The bias shall, however, be clearly stated and recorded along with the uncertainty estimate. If a bias adjustment is made prior to reporting a result (e.g., adjusting for recovery on a sample that is spiked with a known amount of substance), then an additional source of uncertainty is introduced and must be included in the uncertainty estimate. However, if LCS data routinely include adjustments for recovery, then the error from the adjustment is already included in $\mathbf{SD_P}$ and does not need to be added again.

   It is recommended that 20 or more individual LCS data points be obtained to estimate $\mathbf{SD_P}$. The estimate of combined uncertainty is then expanded using the formula:

   MEASUREMENT UNCERTAINTY FOR A DEFINED MATRIX (LCS) = k x SD_P,
   where k (the coverage factor) equals 2 (for 95% confidence)

   If fewer than 20 LCS results are available, the coverage factor should be the appropriate t statistic for 95% confidence for the associated number of degrees of freedom (10=2.228, 20=2.086, 30=2.042, 40=2.021, 60=2.000, 120=1.980 & ∞=1.960, NIST SP260-100: 1993 Table B.3.4).

   NOTE 1: MU estimates from LCS samples should only include data from analysis runs that were determined to be "in control", and should exclude data from runs that were determined to be "out of control" and where reasons for the problem were identified and corrected. When there was no explanation for the "out of control" signal, it might reflect actual uncertainty and should be retained in the MU estimate. However, this depends upon what the result of a root cause investigation revealed, for example if the investigation revealed that the out of control event was not due to an assignable cause.
   NOTE 2: If single LCS results are used in MU calculations but the average of multiple results is reported to the client, then SD_P has to be divided by the square root of the number of measurements used in creating the average.
   NOTE 3: Stated uncertainties for reference materials are usually quite small and are generally considered to be included in the uncertainty calculations for an LCS that is run through all method steps. If reference material uncertainties are significant they should be combined with $\mathbf{SD_P}$ using the root sum square **(RSS)** method.

Method validation data may be used to estimate measurement uncertainty if the validation data were determined by studies that are consistent with ISO 5725. Use of these data also requires that the laboratory has demonstrated its competence with the method, as determined by criteria below.

The laboratory may use a published SD for reproducibility **(SD_R)** as an estimate of combined standard uncertainty under the following conditions:

1. The validation study included all sources of uncertainty (including sample preparation and different analysts)
2. The laboratory has acceptable bias
3. The laboratory has acceptable repeatability, or the estimate is modified appropriately.

To demonstrate competence with a method, the laboratory must calculate the SD for laboratories **(SD_L)**, as the quadratic difference between reproducibility and repeatability **(SD_r)** from the validation study **($SD_L = \sqrt{(SD_R^2 - SD_r^2)}$).** Then the laboratory must estimate their bias using reference materials or other procedures, and estimate their repeatability using a replicates study at the appropriate level.

The laboratory must demonstrate competence with the method by showing that:

1. Their Bias **$< 2SD_L$**
2. Their Repeatability **$< \sqrt{F} \times (SD_r)$**, with F taken from a statistical F table using appropriate degrees of freedom and 95% confidence. The laboratory has an option to use **1.5** as a low limit for **$\sqrt{F}$** (and therefore a tight criterion).

   NOTE: F tables are found in all introductory statistical textbooks and in many computer packages and calculators. Unfortunately the format varies in different presentations regarding the numerator and denominator degrees of freedom and significance level **(α or $\alpha/2$)**. For the purposes of comparing **SD_r** with a lab's repeatability, use the number of observations used to estimate the **SD_r** as the numerator degrees of freedom and the number of replicates used to estimate the laboratory's repeatability as the denominator degrees of freedom. Look only for significance at the low end (repeatability much larger than **SD_r**), so use a one-sided F, with **α=0.05**. As a rough rule, if the repeatability is less than **1.5 times SD_r**, it is acceptable.

If a laboratory has much lower repeatability than **SD_r**, then this lower estimate should be combined with the **SD_L** using **RSS** to obtain a lower estimate of combined uncertainty. Similarly, if a laboratory has acceptable bias, but their repeatability is larger than the criterion, then the laboratory may combine their repeatability with the validation study SD_L to obtain a larger combined uncertainty estimate.

If the validation study did not include all steps in the method, then standard uncertainties from these steps may be added to the **SD_R** with the **RSS** method.

The estimate of combined uncertainty (usually **SD_R**) is then expanded using the following formula:

Measurement Uncertainty for a Defined Matrix = k x SD_R,

where k (the coverage factor) equals 2 (for 95% confidence)

For test methods that need identification of all components of uncertainty and detailed measurement uncertainty budgets, these estimates are to be calculated in accordance with published methods that are consistent with those described in Class III, above

REPORTING MEASUREMENT UNCERTAINTY:

Measurement uncertainty is to be determined for all methods in Categories II and III, and is to be reported when one or more of the following conditions occur:

1. When requested by the client, for example when a result is compared to a decision point.

2. When required by specification or regulation

In these cases, the laboratory must report the expanded uncertainty in the same units as the measurement result and with the same number of significant digits as the reported value. The coverage factor must be included in the uncertainty statement. If the MU was estimated using relative SDs or percentage relative SDs, the percentage must be transformed into the reported units prior to reporting the uncertainty.

If the method has a known bias and this bias was not adjusted (for example, adjustment for recovery), this bias should be reported in addition to the result and the uncertainty.

For example, a measurement method has an average recovery of 89% of the target analyte, and the expanded measurement uncertainty has been estimated as 2.3% at levels below 300ppm. A test result is 210 ppm, and the result is used to prove conformance with a specification limit of 300ppm. The result could be reported as follows:

Sample result = 210 ppm. The expanded uncertainty of this result is +/- 5ppm, with a coverage factor of 95%. This method has an average recovery of 92%, or at this level, a possible bias of 23ppm.

Table 1: Summary of Measurement Uncertainty Calculation Requirements Based on Category of Testing

Category	Classification	Description	UM requirement	Example
Category I	FDA-cleared or approved or Lab Developed Test (LDT) (validated)	Results reported as Pos/Neg (Qualitative)	No MU required.	RPR agglutinin test
Category I	FDA-cleared or approved	Results reported as +1,+2, +3 etc. (Qualitative)	No MU required.	Urobilinogen on a Urinalysis strip
Category II	FDA-cleared or approved	Results reported as a measured quantity (Quantitative)	MU estimate using long-term imprecision (s) with adjustment for bias	Any test on a device that gives a direct read out as a quantity. Qualitative tests based on a quantifiable absorbance cutoff.
Category III	Modified FDA-or approved Test or LDT or Lab Modified Test	Results reported as a measured quantity (Quantitative)	MU estimate using long-term imprecision (s) with adjustment for bias. When this is not possible, MU estimate based on and consistent with Guide to Uncertainty of Measurement (GUM)	Any test developed by the laboratory that expresses the result as a quantity. Any test that is modified by the laboratory that expresses the result as a quantity.

DOCUMENT REVISON HISTORY

Date	Description
04/13/20	➢ Complete rewrite for simplification and consistency with current ISO documents ➢ Integrated into Qualtrax ➢ Updated Header/Footer to current version ➢ Updated format and font for consistency ➢ Added Qualtrax hyperlinks